

## 摘要

通过硬件设计与软件计算的有机结合使得数码相机能够输出颜色外貌符合人眼视觉感知的高质量图像，已成为现阶段数字成像学术研究与生产制造中的共同目标之一。为获得可在不同媒介中进行存储、传输或展示的输出图像，数码相机需要对图像传感器所记录的原始图像数据进行一系列的分析与处理，这一过程通常被称为数字图像信号处理流程（image signal processing pipeline）。典型的图像信号处理流程既需要包含白平衡校正、颜色校正、阶调映射等用于模拟人眼视觉感知特性的图像处理模块，同时还需要包含非线性修正、去噪、非均匀性校正等用于消除设备相关特性以及系统误差的功能性模块。为了获得高质量的输出图像，图像处理算法设计者以及硬件制造者在数码相机出厂前通常需要对其进行一系列的标定与测试，以确保各模块均工作在“最佳状态”。

为解决传统的图像信号处理标定方法中存在的步骤繁琐、流程复杂等问题，本文首先对数码相机生成原始响应信号这一物理过程进行了完整建模，并通过对测试相机的光电转换特性、模拟电路非线性、数字增益、通道串扰等系统特性参数进行估计，构建出与测试相机具有几乎相同响应特性的“虚拟相机”。利用该虚拟相机，我们能够对任意测试场景对应的原始响应值图像进行预测，并以极高的效率为图像信号处理流程中各模块的标定工作提供训练、测试数据以及优化依据。实验结果表明，本文提出的数码相机原始响应值预测模型对于不同的照明条件、不同的拍摄对象以及不同的拍摄参数均达到了令人满意的预测精度。

在原始响应值预测模型的基础上，本文同时对图像信号处理流程中的空间非均匀性校正、白平衡校正、颜色校正以及图像后处理等与数字图像颜色复现密切相关的模块进行了深入分析，并提出了相应的解决方案以改善现有方法与技术中存在的不足。在本文提出的图像信号处理流程设计方案中，首先使用基于拍摄参数的系统误差修正技术对传感器输出图像中的系统噪声进行抑制，并使用光源自适应的二维 B-样条曲面对图像中的亮度以及颜色非均匀性进行补偿。在自动白平衡校正模块中，基于中性色像素筛选的思路，提出了结合色域映射与中性色像素统计的 GMap-NPStat 算法，其工作原理直观可靠，且在性能与模型复杂度之间取得了理想的平衡，十分适合部署于现阶段的移动计算平台之中。针对传统的统计类颜色恒常性算法中存在的一些缺陷，本文还提出了一种基于深度卷积神经网络的颜色恒常性计算模型，该模型在增加少许计算资源的前提下显著提高了光源颜色

估计的准确性，且兼具了较强的泛化能力与可解释性，具有良好的应用前景。完成白平衡校正后，采用光源自适应的根式多项式模型对图像进行颜色校正，以此实现设备相关颜色空间至设备无关颜色空间的跨光源颜色复现。最后，基于 CAT02 色适应模型对图像的整体颜色外貌进行二次调节，以获得更加贴合人眼视觉感知的输出图像。本论文在大量仿真以及实测数据的基础上对所提出的图像信号处理流程进行了分析与评估，结果表明，该方案能够可靠地对大多数拍摄场景进行准确的颜色复现，且同时具备了对于边缘场景的适应能力。

为了进一步论证本论文研究成果的可行性与实用性，我们以自动白平衡校正、颜色校正、同色异谱失配分析以及图像后处理等模块作为范例，对原始响应值预测模型在图像信号处理流程标定任务中所起的实际作用以及潜在用途进行了展示。最终的实验结果表明，利用本论文提出的图像信号处理流程方案，并结合原始响应值预测模型，我们能够在实验室环境下对不同图像处理模块进行端到端的参数优化，由此在不降低数字图像颜色复现精度的前提下显著提高数码相机的标定效率。

**关键词：**图像信号处理流程、响应值预测模型、光谱灵敏度估计、颜色复现、颜色恒常性、自动白平衡、颜色校正、同色异谱、色适应变换

## Abstract

It is becoming a common pursuit for image researchers in both academic and industrial domains to integrate elaborate algorithms development with hardware design and produce high-quality digital images with color appearance coinciding with human visual perception. The image processor in a camera conducts a series of analyses and manipulations to the sensor's output raw signals, which are commonly known as image signal processing pipeline (ISP pipeline), in order to simulate human visual system's features, e.g., color constancy, tone mapping, as well as to get rid of device-dependent characteristics and system errors such as non-linearity, noise and non-uniformity. To produce pleasing output images, it is necessary to perform calibrations for ISP modules before they reach their "optimal status".

The traditional procedures to calibrate ISP modules are overelaborate and time-consuming. To address this issue, in this study, we modeled the physical process of raw responses formation for digital cameras and then constructed a "virtual camera" with practically identical response characteristics as the test camera by estimating its system parameters, for example, spectral sensitivity functions, non-linearity of analog circuit, digital amplification gains and cross-talk effect. With the virtual camera, we were able to accurately predict the camera raw images for any test scenarios, and consequently to generate training and testing data for ISP calibration and optimization. The experimental results indicate that the proposed raw responses prediction model is capable of achieving satisfactory prediction accuracy for different illumination conditions, different objects and different capture parameters.

Based on the raw responses prediction model, we also investigated and proposed novel methodologies and technologies for digital cameras' ISP design. In the proposed ISP pipeline, the system errors involved in the raw image were first corrected by the noise correction module, and the spatial non-uniformity was compensated by a set of reconstructed B-spline surfaces in an illuminant-adaptive fashion. To reproduce objects' intrinsic colors independent of the light source, an automatic white balancing algorithm, dubbed GMap-NPStat, was presented to adjust the color appearance of the image, which worked by incorporating neutral-pixel-counting method and gamut-mapping method. The flexibility and viability make GMap-NPStat particularly suitable to be deployed on the mobile platform. In addition, an interpretable convolutional-neural-network based model for computational color constancy was also proposed, which significantly improved the accuracy of illuminant color estimation at the cost of slightly increasing computational complexity. After white balancing, the optimized root-polynomial color correction model was

employed to convert device-dependent responses under the test illuminant to the device-independent colorimetric values under the canonical illuminant. Finally, the overall color appearance of the image was adjusted again by the post-processing module based on the CAT02 chromatic adaptation transform model, which made it appear more natural and preferable as the human visual system perceived. The evaluation on both synthetic and real data suggests that the proposed ISP pipeline achieves competitive color reproduction accuracy for most test scenarios and also exhibits robustness for the corner cases.

To further validate the feasibility of the proposed solution, several examples were presented to illustrate the practical and potential usages of the raw responses prediction model in ISP design and calibration. The results prove that, with the help of raw responses prediction model, the end-to-end parameters optimization for image processing modules could be handled in the laboratory environment, which greatly improves the efficiency of ISP calibration whilst no performance degradation should be concerned.

**Keywords:** image signal processing pipeline, responses prediction model, spectral sensitivity functions estimation, color reproduction, color constancy, automatic white balancing, color correction, metamer mismatching, chromatic adaptation transform

## 主要符号说明

$a$ 或 $A$	标量
$\mathbf{a}$ 或 $\mathbf{A}$	向量、矩阵或张量
$P(\lambda)$	光谱功率分布函数
$s_k(\lambda)$	数码相机 $k$ 通道的光谱灵敏度函数， $k \in \{r, g, b\}$
$X, Y, Z$ 或 $\mathbf{X}$	CIEXYZ 三刺激值或刺激值向量，且有 $\mathbf{X} = [X, Y, Z]$
$D_k$ 或 $\mathbf{D}$	$k$ 通道数字响应值或响应值向量，且有 $\mathbf{D} = [D_r, D_g, D_b]$
$\mathcal{I}$	图像
$\mathcal{G}$	白平衡校正增益系数
$\mathbf{C}$	CMOS 通道串扰矩阵
$\mathbf{H}$	二维直方图矩阵
$\mathbf{I}$	单位矩阵
$[\dots]$	向量或矩阵
$\{\dots\}$	集合
$(\cdot)^\top$	向量或矩阵转置
$\ \cdot\ _p$	$\ell_p$ 范数



# 目录

致谢 .....	I
摘要 .....	III
ABSTRACT .....	V
主要符号说明 .....	VII
<b>第一章 绪论 .....</b>	<b>1</b>
1.1 研究背景及意义 .....	1
1.1.1 研究背景 .....	1
1.1.2 研究目的及意义 .....	2
1.2 国内外研究现状 .....	3
1.2.1 数码相机特征化技术 .....	4
1.2.2 图像信号处理流程中的常见算法 .....	6
1.2.3 同色异谱 .....	9
1.3 本论文主要研究内容 .....	10
<b>第二章 颜色科学与影像技术理论基础 .....</b>	<b>13</b>
2.1 辐射度学与光度学基础知识 .....	13
2.1.1 辐射量 .....	13
2.1.2 光谱辐射量 .....	14
2.1.3 光度量 .....	15
2.1.4 辐射量与光度量之间的转换 .....	17
2.2 色度学基础知识 .....	18
2.2.1 CIE 标准色度系统 .....	18
2.2.2 均匀颜色空间与色差评价 .....	21
2.2.3 光源与色温 .....	23
2.3 色貌现象与色适应模型 .....	26
2.4 本章小结 .....	28
<b>第三章 数码相机原始响应值预测模型 .....</b>	<b>29</b>
3.1 数码相机原始响应值构成模型 .....	29

3.1.1 感光单元表面光谱辐照度构成模型 .....	30
3.1.2 电荷信号构成模型 .....	32
3.1.3 数字信号构成模型 .....	33
3.2 数码相机的系统误差修正 .....	36
3.2.1 图像传感器的噪声模型 .....	37
3.2.2 数码相机系统误差的标定与修正 .....	40
3.3 基于光谱灵敏度函数估计的数码相机原始响应值预测模型 .....	48
3.3.1 原始响应值构成模型中的非线性效应 .....	50
3.3.2 通道串扰效应 .....	52
3.3.3 原始响应值构成模型中的参数估计 .....	54
3.3.4 实验设置 .....	64
3.3.5 结果与分析 .....	67
3.4 本章小结 .....	69
<b>第四章 数码相机图像信号处理算法研究 .....</b>	<b>71</b>
4.1 图像信号处理流程概述 .....	71
4.2 空间非均匀性校正 .....	72
4.2.1 亮度非均匀性与颜色非均匀性 .....	73
4.2.2 基于 B-样条曲面的空间非均匀性校正算法 .....	77
4.3 颜色恒常性与自动白平衡校正 .....	84
4.3.1 结合色域映射与中性色像素统计的自动白平衡算法（GMap-NPStat） .....	87
4.3.2 基于特征图二次卷积的自动白平衡算法（ReWNet） .....	105
4.4 颜色校正 .....	119
4.4.1 基于根式多项式矩阵的颜色校正算法 .....	121
4.4.2 物体色同色异谱失配与颜色失真 .....	127
4.5 图像后处理 .....	140
4.5.1 基于色适应变换模型的图像颜色外貌调节算法 .....	140
4.6 本章小结 .....	143
<b>第五章 原始响应值预测模型应用实例 .....</b>	<b>145</b>
5.1 原始响应值预测模型在自动白平衡校正模块参数标定中的应用 .....	145

5.1.1 相机 RGB 颜色空间至正交色度平面的变换.....	145
5.1.2 NPStat 算法中中性色区域的预设定 .....	148
5.1.3 二维色域映射算法中标准色域的计算.....	150
5.2 原始响应值预测模型在颜色校正模块参数标定中的应用 .....	151
5.2.1 二次项根式多项式颜色校正矩阵的标定 .....	151
5.3 原始响应值预测模型在图像后处理模块参数标定中的应用 .....	153
5.3.1 拍摄场景中白场亮度的估计 .....	153
<b>第六章 总结与展望.....</b>	<b>155</b>
6.1 研究工作与创新点总结 .....	155
6.2 存在的不足与未来工作展望 .....	156
参考文献 .....	159
附录一 .....	175
附录二 .....	176
插图及附表清单 .....	177
作者简介与主要研究成果 .....	185



# 第一章 绪论

## 1.1 研究背景及意义

### 1.1.1 研究背景

图像是现代人们认识、探索世界的最重要的信息载体之一。随着技术的进步与发展，人们获取图像信息的方式也在不断变化，各种成像设备应运而生。以 20 世纪 50 年代数字录像机的量产为标志，电子成像技术成为了传统的光化学成像技术之外的一种重要的图像获取方式。相比于传统相机感光底片的光化学成像，电子成像具有使用成本低、数据存储灵活、成像周期短、可定制程度高、外形轻便等诸多优点。从上世纪 80 年代开始，得益于高感光度半导体材料研发水平和数字存储技术的成熟，电子成像技术发展迅速，并广泛应用于航空航天、医疗、公共安防、新闻传媒等领域<sup>[1-3]</sup>。在各类电子成像设备中，以电荷耦合感光器件（charge coupled device, CCD）及互补金属氧化物半导体（complementary metal-oxide-semiconductor, CMOS）为核心的彩色数码相机无疑是使用最为广泛、与工业生产和日常生活关系最为密切的一类设备。

在彩色数码相机普及伊始，如何提高空间分辨率以及如何加快数据存储速度是该领域内的主要研究目标，而近年来随着相关硬件技术的日臻成熟，关于图像质量的优化与提升正逐渐成为数字信号处理领域内的热门方向<sup>[1]</sup>，相关的研究分支也日益增多<sup>[4-6]</sup>。数字图像质量的优化可以分为以人眼视觉系统感知为导向的图像真实复现（image reproduction）以及以观察者主观喜好性为导向的图像增强（image enhancement）两大类，其中图像增强通常需要以真实复现的图像作为基础，并在此之上进行喜好性调节以获得更佳的主观图像质量<sup>[7-10]</sup>。

在所有可能影响图像复现真实性的因素当中，场景中的颜色信息还原准确性无疑是用户能够最直观感知的指标之一。为了使数字图像的颜色外貌与用户直接观察拍摄对象时所感知的颜色外貌尽可能一致，数码相机在其内部通常需要对采集到的原始图像信号进行一系列的处理，以弥补相机感光特性与人眼视觉感知特性间的差异。针对颜色复现准确性这一指标，现有的数码相机图像信号处理流程中仍然存在一些不足，例如，自动白平衡模块尚无法对所有拍摄场景进行准确的白平衡校正<sup>[11]</sup>、光源色同色异谱现象可能导致颜色校正模块出现误判<sup>[12]</sup>、某些特殊场景的输出图像外貌与人眼视觉感知明显不符，等等。这些问题都阻碍了图像真实复现这一目标的实现。因此，如何改善数字图像颜色复现的精度并提

升最终面向用户的输出图像质量，成为了现阶段学术界和工业界共同的研究目标。

此外，不同的相机模组往往具有不同的输入输出特性，故为使不同设备的输出图像具有一致的颜色外貌，数码相机制造商在设备出厂前通常需要对其进行一系列的标定与测试。目前，在主流的数码相机标定流程中，实验者需要在各种复杂场景下采集大量的测试数据，从而为确定各模块标定参数提供足够的依据。这种传统的标定方法步骤繁琐，流程复杂，且需要针对不同的模块单元重复多次实验，既增加了购置标定设备的成本，也耗费了大量的人力物力。因此，如何简化标定流程、提高标定效率，也成为了数码相机制造者最为关注的问题之一。

### 1.1.2 研究目的及意义

提高数码相机颜色复现的准确性并简化其标定流程，在研究与生产中都具有重要意义。在计算机视觉领域中，生物特征识别<sup>[13]</sup>、图像分类与检索<sup>[14, 15]</sup>、三维重构<sup>[16, 17]</sup>、图像恢复<sup>[18]</sup>等应用都需要使用大量具有正确颜色外貌的图像对模型进行训练，因此训练图像的颜色复现准确性将直接对计算机视觉模型的最终性能产生影响；在生产生活中，基于机器视觉的产线自动化控制、基于成像的测量和分析、移动终端的图像拍摄等应用也都对图像的颜色复现准确性提出了一定的要求。

利用标准测试场景或测试样本对数码相机系统参数进行标定时，需要对噪声修正、暗电流校正、空间非均匀性校正、自动白平衡校正以及颜色校正等模块进行大量的数据采集、分析及调试工作。这一过程不仅费时费力，也容易引入一些难以预估的实验误差。例如，光源的辐射光谱随点亮时间的增加而出现漂移、长时间拍摄致芯片发热从而改变传感器的噪声特性、多次拍摄的过程中引入人为误差等等。除了存在不可避免的实验不确定性之外，这种传统的标定过程更明显的缺陷在于，由于实际拍摄时所能囊括的场景和样本的数量总是十分有限的，因此标定与测试很难覆盖到用户实际使用设备时所可能遇到的各种情况。这一问题在评估自动白平衡算法和颜色校正算法时显得尤为突出。

本论文在目前主流的图像信号处理流程标定方法的基础之上，对如何提高空间非均匀性校正、白平衡校正、颜色校正、图像后处理等模块的性能进行了探索，并提出了基于相机光谱灵敏度函数估计的数码相机原始响应值预测模型，建立起目标场景中辐射量与相机数字响应值之间准确的映射关系。该模型允许用户通过仿真的手段重构出拍摄场景对应的数字图像信息，并为图像信号处理流程中各模块的参数调试过程提供参考，从而大幅提高数码相机的标定效率，实现标定流程的自动化与批量化。此外，将该响应值预测模型与高

光谱图像（hyperspectral images）相配合，研究者可以方便地对数码相机实际使用过程中可能遇到的各种场景进行模拟，从而为自动白平衡、颜色校正等算法的设计与优化提供翔实可靠的依据。

数码相机图像信号处理流程的优化作为与生产应用高度相关的一个研究课题，目前在学术研究与工业生产中已逐渐呈现出各自为营的趋势。一方面，两者的研究目标并不相同：学术研究领域内提出的图像处理算法，往往更加注重性能与可解释性，而对模型的计算复杂度、资源开销等指标并不敏感；而工业生产中，研究者在设计图像信号处理算法时通常更加关心运行效率、性价比、鲁棒性等指标，往往需要为了满足实时性和低功耗的要求而在性能上做出妥协<sup>[19]</sup>。另一方面，学术领域中的一些研究工作往往只针对单一功能的图像处理模块进行探索，并不涉及完整的信号流处理过程；而数码相机制造商出于保护知识产权的需要，又通常对相机内部与图像处理算法有关的具体技术细节进行严格保密。这一矛盾导致了现阶段对数码相机图像信号处理及其标定流程进行完整介绍的公开资料寥寥无几。针对这一现象，本论文拟将同时从学术研究与工程设计两个角度对数码相机图像信号处理流程中与“颜色准确复现”有关的方法与技术进行深入探究，在提出面向应用层面的图像处理算法的同时，也对现阶段数字图像颜色复现技术中存在的一些不足之处进行分析并提出改进思路。

## 1.2 国内外研究现状

针对数字图像颜色准确复现这一目标，不同的应用场景对于算法的设计及标定实验的实施有着不同的要求。在摄影测量学（photogrammetry）与摄影光度学（photographic photometry）领域，数码相机往往被视为测量仪器，研究者通过对图像进行分析从而对拍摄场景中的辐射度或光度信息进行估计。这一类应用对于数字图像颜色复现精度的要求较高，但允许用户在线标定，且不要求模型具有很强的泛化能力。在计算摄影学（computational photography）领域，数码相机往往被视为一种视觉信息记录工具，其输出应尽可能真实地还原拍摄者在拍摄场景中的视觉感知。由于人眼的视觉感知带有很强的主观性质，这一类应用不需要对场景的物理参数进行精确复现，但对于图像的真实性（reality）及喜好性（preference）有着较高的要求，且往往需要模型具有很强的泛化能力。

针对以上两类应用，本节将主要从数码相机的特征化技术及图像信号处理流程两个角度对现有的研究成果进行简述，其中前者旨在追求对拍摄对象的客观辐射度或光度、色度

信息进行准确复现，而后者则通过一系列的信号处理操作使得输出图像与用户的主观视觉感知尽可能贴合。此外，对于图像捕获过程中存在的同色异谱现象及其对于颜色复现产生的影响，本节中也将给出概要性描述。

### 1.2.1 数码相机特征化技术

根据研究对象的不同，数码相机特征化技术可分为光谱特征化(spectral characterization)与色度特征化(colorimetric characterization)两类。光谱特征化是对数码相机中光信号转变为数字信号的完整物理过程进行建模，也是对其固有特性最彻底的描述。然而光谱信息获取难度较大，且多数应用并不需要从光谱层面对数码相机进行标定，因此在实际应用场景中色度特征化技术受到了更为广泛的关注。

#### 1.2.1.1 数码相机光谱特征化技术

以 RGB 数码相机为例，其图像传感器上某一感光单元(像素)输出的原始数字响应值  $D_k$  与光信号  $\Phi(\lambda)$  之间，可通过如下积分模型进行定性描述：

$$D_k = T \int_{\Omega} \Phi(\lambda) q_k(\lambda) d\lambda, \quad (1.1)$$

式中  $k \in \{r, g, b\}$  表示该感光单元所对应的颜色通道， $T$  表示积分时间， $\lambda$  表示波长， $\Omega$  表示有效的波长响应范围， $\Phi(\lambda)$  表示该感光单元表面接收的光谱辐射通量(spectral radiant flux，详见第 2.1.2 小节)， $q_k(\lambda)$  表示  $k$  通道的光电转换效率函数。显然，如果能够准确地获知一台成像设备的光电转换效率函数  $q_k(\lambda)$ ，那么对于任何被摄场景，我们都能利用式(1.1)建立起物理辐射量与数字量之间的转换关系。

目前主流的获取数码相机光谱灵敏度函数的方法包括物理测量法与参数估计法两类。物理测量法使用单色仪或窄带滤光片生成各波长上的准单色光，通过对相机关于各单色光的响应值强度进行分析从而确定其光谱灵敏度函数。为了确保光源的空间均匀性，实验者通常需要利用光纤将准单色光耦合进积分球内，经漫反射后再到达各个感光单元表面<sup>[20]</sup>。物理测量法直观、易操作，但由于测量过程中存在诸多系统误差，利用该方法得到的光谱灵敏度数据未必能够准确地反映成像系统的光谱响应特性<sup>[21]</sup>。

相比于物理测量法，基于标定过程的参数估计法能够更加可靠地对成像系统的光谱灵敏度函数进行重构，且无需使用昂贵的单色仪与积分球。Barnard 等人提出了基于带约束非负最小二乘法的光谱灵敏度函数估计方法，并用指数模型拟合了预测响应值与真实响应值之间的非线性关系，最终得到 1% 量级的响应值估计误差<sup>[22]</sup>。Jiang 等人利用主成分分析

法对不同相机光谱灵敏度函数之间的关系进行了分析，并分别采用傅里叶函数、多项式函数、径向基核函数作为基底对每台测试相机的光谱灵敏度函数进行了表征<sup>[23]</sup>。Prasad 等人认为相机的光谱灵敏度函数与标准观察者色匹配函数之间应尽可能满足 Luther-Ives 关系，因此在最小二乘法的基础上进行了改进，将两组函数之间的线性相关程度作为优化目标<sup>[24]</sup>。Darrodi 等人则对光谱灵敏度函数估计问题进行了综述性介绍，并分析了光源光谱功率分布函数以及训练色块的选取对估计精度可能产生的影响<sup>[25]</sup>。

在已有的文献之中，研究者通常采用两种方式对光谱灵敏度函数的估计精度进行评估。第一种方式以估计出的光谱灵敏度函数与实测值之间的光谱差异作为估计精度的评价指标，第二种方式则将估计出的光谱灵敏度函数与实测数据同时代入积分模型（1.1），并在积分后响应值的基础上判断两者间的误差。对于使用光谱灵敏度函数进行响应值预测的应用来说，无疑是使用第二种方式得到的评估结果更具说服力<sup>[26]</sup>。

### 1.2.1.2 数码相机色度特征化技术

如前文所述，不同的数码相机往往具有不同的光谱响应特性，因此即使对于同一拍摄对象，不同相机也可能输出不同的原始数字响应值。色度特征化技术通过建立数字响应值与色度刺激值之间的映射关系，实现了数字信号从设备相关颜色空间至设备无关颜色空间的变换\*。

对于大部分数码相机来说，其光谱灵敏度函数与标准观察者色匹配函数之间并不满足线性变换条件，因此积分后的数字响应值与色度刺激值之间也无法使用简单的线性模型进行匹配。现有的色度特征化模型，例如多项式拟合<sup>[30,31]</sup>、分区域函数<sup>[27,32]</sup>、神经网络<sup>[33]</sup>等，往往通过在模型中引入一定的非线性从而达到更高的匹配精度。然而，这些经验模型的性能通常高度依赖于成像设备本身的输入输出特性，且泛化能力较弱，一旦测试阶段与训练阶段的光环境发生变化，特征化精度也将出现较明显的下降。

此外，对于一些不适合使用上述模型的设备，通常也会采用查表法直接建立起数字响应值与色度刺激值之间的映射。查表法通过对设备相关颜色空间以及设备无关颜色空间分别进行三维网格划分从而建立其两者子空间之间的查找对应关系。显然，为达到较高的匹配精度，查表法需要大量的基础数据，且建表效率较低；当光环境发生变化时，表的更新

---

\* 在一些文献中，数码相机的色度特征化也指计算数字驱动值至色度值之间的映射<sup>[27-29]</sup>，其中数字驱动值表示经过机内一系列处理流程后最终输出图像中的像素数值。本论文将严格区分数码相机的“原始响应值”与“输出驱动值”，若无特殊说明，文中的“响应值”均指图像传感器直接输出的或位于图像信号处理流程中的线性响应信号。

也较为繁琐。

### 1.2.2 图像信号处理流程中的常见算法

特征化技术旨在对相机固有的设备特性进行描述，而图像信号处理流程则通过一系列的图像数据分析与处理操作，使得数码相机能够获得与人眼视觉系统一致甚至优于人眼视觉系统的图像信息采集能力。

对于现阶段的消费级数码相机而言，一套基于彩色滤色片阵列输出图像（color filter array image，后文中简称为 CFA 图像）的完整图像信号处理流程（image signal processing pipeline，后文中简称为 ISP 流程）通常包含以下模块：非线性校正、暗电流校正、系统噪声修正、反马赛克、空间非均匀性校正、白平衡校正、颜色校正、颜色空间变换、降噪、图像增强、编码存储。同时，当用户开启自动曝光和自动对焦选项时，ISP 还需要根据图像内容实时地对光圈、快门、对焦马达等机械结构进行反馈控制，以确保图像得到合适的曝光和对焦。

本小节仅针对数字图像颜色准确复现这一目标，对典型 ISP 流程中空间非均匀性校正、自动白平衡校正以及颜色校正三个模块的研究现状进行简述。

#### 1.2.2.1 空间非均匀性校正

数字成像系统中的空间非均匀性（spatial nonuniformity/inhomogeneity）指的是具有相同辐亮度的物体成像于图像传感器不同位置时对感光单元表面产生不同辐照度这一特性，在摄影领域中也被称为镜头阴影（lens shading）或渐晕（vignetting）。导致空间非均匀性的成因众多（例如，光学镜头存在像差和色差、传感器各处对应的主光线具有不同入射角、光阑口径有限等），往往只能通过光线追迹的方式进行仿真，难以使用数学模型直接表达。在实际的 ISP 流程中，一般采用“先标定后调用”的方法进行校正，即在出厂前先对若干组非均匀性曲面进行拟合，并将拟合参数存储于相机的内置存储空间中，待使用时再根据实际情况从中选取调用。

Münzenmayer 等人为改善图像中的亮度非均匀性，将 RGB 图像转换至 YUV 颜色空间中，并对低通滤波后的 Y 分量逐像素计算校正系数<sup>[34]</sup>。该方法的局限性在于其只能对相机系统中的亮度非均匀性进行修正，而无法解决颜色非均匀性。Young 采用二次曲面对显微成像系统中的空间非均匀性曲面进行拟合<sup>[35]</sup>，该模型适合空间非均匀性曲面较为平滑的相机系统，但对于某些具有不规则衰减曲面的设备，并不能得到理想的拟合精度。Yu 采用与 Young 类似的方法进行非均匀性校正曲面的标定<sup>[36]</sup>，但将二次曲面改为查找表和二维双

曲函数，并在拟合前首先使用小波降噪对曲面进行平滑处理。Tajbakhsh 则采用双线性样条曲面对非均匀性曲面进行拟合<sup>[37]</sup>。由于样条曲面允许对待标定区域进行分区处理，因此该方法能够较好地适应各种不规整的衰减曲面，具有较好的通用性。

除了“先标定后调用”的方法之外，也有许多基于单幅图像的非均匀性校正（single-image shading correction）算法相继被提出。Likar 等人通过最小化图像熵的方法对非均匀性校正曲面中的参数进行估计<sup>[38]</sup>。Fuentes 等人则在该方法的基础上对图像熵的定义以及直方图变换方法均进行了改进，从而缓解了优化过程中各参数陷入局部最优的问题<sup>[39]</sup>。Zheng 等人通过对图像中各像素的半圆切向梯度（semicircular tangential gradients）以及径向梯度（radial gradients）进行统计分析，发现这两类梯度的对称性与相机系统的非均匀性之间存在高度相关性，并提出了基于径向梯度最小化的图像的非均匀性校正算法<sup>[40]</sup>。由于单幅图像非均匀性校正算法往往基于一些先验假设，一旦在实际场景中这些假设无法得到满足，则校正性能可能出现明显的下降，例如，当场景中存在着由于照明不均匀性而引发的亮度不均匀性时，此类算法将无法有效区分非均匀性的来源。此外，当图像中非均匀性成分的频谱与实际拍摄对象的频谱存在重叠时，此类算法也将大概率失效。

### 1.2.2.2 自动白平衡校正

白平衡校正是实现数字图像颜色准确复现过程中的关键步骤之一，也是颜色恒常性计算（computational color constancy）概念中的核心内容。

根据是否需要提供带标签的训练数据，自动白平衡校正算法可被分为基于先验假设的算法以及基于学习的算法两大类。基于先验假设的自动白平衡算法通常认为，当图像的白平衡得到正确的校正之后，其颜色分布将满足某些特殊的规律，因而可以以此为参照对图像三通道的增益系数进行调节。灰度世界算法（Gray World）与白板算法（White Patch）是最为基础的两种基于先验假设的自动白平衡算法，前者假设某一场景中所有物体表面的光谱反射比均值对应于一条中性的反射曲线，故对所有像素的响应值计算均值之后，其色品应呈现出中性性质<sup>[41]</sup>，而后者则认为该规律仅在图像中最亮的部分区域内成立<sup>[42]</sup>。在灰度世界和白板算法的基础上还出现各种算法变体，例如边缘灰度世界算法<sup>[43-45]</sup>、中性像素统计算法<sup>[46-49]</sup>等。Gao 等人使用图像块中的最大值对该区块内所有像素进行归一化，得到局部归一化的反射比估计（LNRE）<sup>[50]</sup>，在对大量图像进行测试后，他们发现无论对于室内或是室外场景，当图像的白平衡参数设置正确时，全局反射比之和与全局 LNRE 之和的比值对于三通道均为常数，故可以利用该性质对图像进行自动白平衡校正。先验假设类自

动白平衡算法通常具有模型简单、执行效率高的优点，但是对于某些不满足先验假设的场景，其性能有可能出现急剧下降，因此并不适合独立地运行于 ISP 之中。

基于学习的自动白平衡算法通常会在大量带标签的训练数据集的基础上对模型中的各个参数进行学习。Forsyth 认为不同光源下由实际物体色构成的色域总是有限的，并且可通过计算得到未知光源至参考光源的映射关系<sup>[51]</sup>，故该算法被称为色域映射算法。在此基础上，各种算法变体也相继被提出<sup>[52, 53]</sup>。除此之外，颜色相关性算法<sup>[54]</sup>和基于小波变换的算法<sup>[55]</sup>也都被证实具有良好的应用性能。近年来，随着机器学习的再度兴起，支持向量机、人工神经网络、深度学习等机器学习方法也被广泛应用于自动白平衡校正算法的设计之中。Funt 等人利用支持向量回归模型对图像的直方图进行分析，并利用交叉检验的方法确定最终的白平衡校正参数<sup>[56]</sup>。Cheng 等人将图像的平均色品、最亮区域色品、RGB 直方图众数色品以及彩度平面众数色品这四个特征作为参数输入决策树中，并利用 ColorChecker 图像数据库<sup>[57]</sup>对决策树中的分支判据进行训练，从而得到最终的自动白平衡校正模型<sup>[58]</sup>。Bianco 等人设计了包含两层卷积层以及两层全连接层的卷积神经网络，并在 ColorChecker 图像数据库上进行训练，得到了当时性能最优的自动白平衡校正模型<sup>[59]</sup>。

相比于先验假设类算法，学习类自动白平衡算法在性能和泛化能力上有所提升，但是对训练数据的数量与质量有着较高的要求，且训练过程费时费力。针对这一问题，McDonagh 等人基于元学习（meta-learning）的概念提出了一种相机自适应的模型迁移方法<sup>[60]</sup>。该方法通过寻找一组对于梯度更新最为敏感的模型初值并在此基础上对待测相机进行训练，从而在极少数训练样本的情况下实现了参数收敛，显著提高了训练效率。如何对自动白平衡校正模型的训练过程进行简化，在模型性能与标定效率之间达到最佳的平衡，也成为了该领域内的一个热门研究方向。

此外，Zapryanov 等人<sup>[61]</sup>与 Hordley<sup>[11]</sup>等人对自动白平衡校正这一研究领域进行了综述性介绍，同时也对各种算法的性能进行了详尽的比较。

### 1.2.2.3 颜色校正

除了白平衡校正之外，颜色校正同样也是实现数字图像颜色准确复现以及广义颜色恒常性过程中必不可少的步骤。与色度特征化类似，颜色校正的目的同样是实现设备相关信息（相机原始响应值）至设备无关信号（XYZ 色度刺激值或 sRGB 驱动值）的映射。两者区别在于，色度特征化过程中的输入值与参考值是基于同样的拍摄环境获得的，而颜色校正过程中的参考值往往是特定环境下的理想颜色信号。因此，色度特征化模型反映的仅

仅是相机本身的特性，而颜色校正模型则隐含了如何让相机在不同照明环境之间进行“适应”这一层意味。

Finlayson 等人认为颜色校正的过程不应改变图像中中性色 (neutral color) 的色品，否则将会破坏白平衡校正的结果。在此基础上，他们提出了使用带白点约束的最小二乘法对颜色校正矩阵进行计算<sup>[62]</sup>。Hong 等人使用多项式模型进行颜色校正，相比于传统的线性模型，多项式模型能够显著提升颜色校正的精度<sup>[30]</sup>。针对多项式模型不具备曝光不变性 (exposure invariant) 这一缺陷，Finlayson 等人提出了根号多项式颜色校正模型<sup>[63]</sup>，该模型允许用户在不同拍摄参数下进行训练和测试。Cherdhirunkorn 等人利用相机 RGB 图像对场景中的光谱信息进行重构，从而实现设备无关、观察条件无关的颜色信息采集，再通过将该光谱信息转换为参考光源下的 RGB 图像从而实现颜色校正<sup>[64]</sup>。Marguier 等人提出了针对肤色的颜色校正方法，对于包含肤色的图像，该模型自动提取出肤色像素，并以此作为颜色校正矩阵计算依据，并将该矩阵应用于整幅图像<sup>[65]</sup>。Zhang 等人则对色差进行了约束，允许经校正后的响应值与参考值在饱和度和明度上存在较大的误差，以此换取色调上的还原精度<sup>[66]</sup>。

虽然各种性能优越的颜色校正模型已被相继提出，但是如何根据拍摄场景中的照明光源自适应选取合适的校正模型这一问题仍然没有得到很好的解决。

### 1.2.3 同色异谱

根据人眼视觉系统中的颜色感知机理可知，两组具有不同光谱功率分布的光信号可能对应一组相同的色度信号，因此这种现象也被称为同色异谱 (metamerism)<sup>[67]</sup>。倘若在 A 观察条件下具有相同颜色外貌的一组物体被置于 B 观察条件下时其颜色外貌不再匹配，则称该现象为同色异谱失配 (metamer mismatching)<sup>[68]</sup>。由于三色数码相机所能获取的光谱信息远少于表征一个物体的反射特性所需要的光谱信息，因此在图像捕获及处理的过程中，同色异谱失配现象将不可避免地影响颜色复现准确性<sup>[69]</sup>。换言之，由于同色异谱失配现象的存在，三色相机理论上便不可能实现完美的跨光源颜色复现。

目前，针对人眼视觉系统的同色异谱失配现象已经得到了比较深入的研究<sup>[70-72]</sup>，但是对于数字成像系统来说，同色异谱失配现象仍具有很高的研究价值<sup>[73]</sup>。在计算机视觉领域中，由于同色异谱失配而导致的颜色复现精度下降，对于图像分割<sup>[74]</sup>、模式识别<sup>[75]</sup>、颜色校正等应用都将产生一定程度的影响。在 Logvinenko 等人关于同色异谱与颜色恒常性的实验中，较高的同色异谱失配系数以及较差的视觉匹配结果均表明，同色异谱对于人眼颜

色恒常性的影响比预期要大得多<sup>[71]</sup>。然而, Zhang 等人使用了超过 2000 万条真实存在的光谱反射比数据对实际拍摄条件下的同色异谱失配现象进行了模拟, 结果表明, 现实情况中的同色异谱失配对数字成像系统带来的影响要远小于理论分析结果<sup>[76]</sup>。

数字成像系统的同色异谱失配程度主要由其光谱灵敏度函数所决定。Hung 等人提出的成像设备显色指数指标建立起了相机同色异谱程度与视觉感知误差间的联系<sup>[77]</sup>。此外, ISO 17312 标准中还提出了数码相机灵敏度同色异谱系数 (DSCSMI) 指标<sup>[78]</sup>, 该指标仿照 CIE 显色指数的计算方法, 通过对 8 个典型样本的色差进行分析从而确定相机的同色异谱失配程度。

### 1.3 本论文主要研究内容

围绕着改进数码相机标定流程以及提升颜色复现精度这两个目标, 本论文基于数码相机原始响应值预测模型, 对数码相机图像信号处理 (ISP) 流程中的方法与技术进行了深入探究。

第一章首先对课题研究背景与研究意义进行了探讨, 并在此基础上确定了本论文的研究目标, 即对现有的数码相机 ISP 流程中与颜色复现有关的方法与技术进行改进, 同时简化 ISP 流程中各模块的参数标定过程。同时, 我们对近年来国内外关于数码相机标定技术以及 ISP 算法开发的研究现状进行了整理与总结, 并分析了其中的不足之处与可改进的方向。

第二章主要介绍了与本论文研究内容高度相关的辐射度学、光度色度学基础知识。当我们将图像颜色外貌与人眼视觉感知之间的一致性作为图像复现结果的评价指标时, 往往还需要考虑照明条件、观察条件、背景环境等光度色度参数之外的实验变量, 因此本章中也对用于描述与预测各种视觉现象的色貌模型进行了简要的介绍。

第三章首先对数码相机的原始响应值构成模型进行了详细说明。根据信号载体的不同, 数码相机原始响应值构成的过程可分为三个阶段: 第一阶段中相机镜头模组通过光学成像的方式收集场景内的辐射能量, 从而将光信号从三维空间传递至传感器平面; 第二阶段中传感器上的感光元件利用光电转换效应完成光信号至电荷信号的转换; 第三阶段中相机内的读出电路将该电荷信号转换为数字信号, 并以彩色滤色片阵列图像 (CFA image) 的形式将其输出, 供 ISP 后续流程调用。为了对该响应值的构成过程进行更加准确的描述, 本章还基于 CMOS 传感器的硬件特性对数码相机的系统误差修正方法进行了介绍, 通过对

CFA 图像生成过程中的各种噪声进行标定与修正，可以对原始数字响应值进行更加准确的预测。最后，阐述了数码相机原始响应值构成模型中的参数估计方法。该方法通过对一定数量的训练样本进行拍摄，能够在无需任何标定仪器辅助的情况下完成数码相机光谱响应特性、系统非线性、数字增益、通道串扰等系统特性参数的估计。

第四章分别对 ISP 流程中的空间非均匀性校正、白平衡校正、颜色校正以及图像后处理等与颜色复现紧密相关的图像处理模块进行了介绍。在空间非均匀性校正一节中，将非均匀性补偿曲面的生成过程与光源颜色估计模块进行配合，从而实现不同光源下的空间非均匀性自适应校正。在颜色恒常性与自动白平衡校正一节中，分别介绍了基于先验假设的光源颜色估计算法以及基于深度学习的光源颜色估计算法。两类算法均具有很强的可解释性，且在模型性能与复杂度之间取得了良好的平衡，十分适合部署于移动平台之中。在颜色校正一节中，使用根式多项式模型实现了设备相关至设备无关、未知光源至参考光源的数字响应值映射，同时基于大量的真实数据对数码相机颜色复现过程中的物体色同色异谱失配现象进行了分析。在图像后处理一节中，提出了一种基于 CAT02 色适应变换模型的图像颜色外貌调节算法，该算法能够根据场景的亮度信息以及光源颜色信息自适应地对图像的整体颜色外貌进行调节，从而使得输出图像具有与观察者实际视觉感知更加贴近的颜色外貌。

第五章以 Nikon D3x 单反相机作为范例，探讨了数码相机原始响应值预测模型在自动白平衡校正、颜色校正、同色异谱失配分析以及图像后处理模块中的应用。借助原始响应值预测模型，我们可通过仿真的手段对任意辐射体的原始数字响应值进行预测，从而“复制”了一台与原相机具有几乎相同响应特性的虚拟相机。该虚拟相机能够以极高的效率为 ISP 流程中各模块的标定实验提供大量的训练数据以及优化依据。

第六章则对本论文的主要研究内容与创新之处进行了总结，由此对其中存在的不足之处进行了分析，并对后续的研究工作予以展望。



## 第二章 颜色科学与影像技术理论基础

### 2.1 辐射度学与光度学基础知识

光是人们最熟悉的“物质”之一，也是最重要的信息载体之一。光信号传播的本质是电磁波辐射。由于人眼只对电磁波辐射中很小一段波长范围内的能量敏感，因此当我们研究关于光辐射能量的计量问题时，不得不把可见光与非可见光区分开来。辐射度学（radiometry）是一门研究 X 射线、紫外辐射、可见光、红外辐射以及其它电磁辐射能量的计量学科，而光度学（photometry）则是一门把可见光的能量与人眼对它的接收特性结合起来进行研究的计量学科<sup>[79]</sup>。由于人眼对于可见光波段内辐射能量的光谱光视效率函数已经建立<sup>[80]</sup>，光度量可以从相应的光谱辐射量中计算出来。因此，广义地讲，目前光度学实质上已是辐射度学的一个组成部分。辐射度学和光度学两者在研究方法上和概念上基本相同，且前者中的客观物理量与后者中的主观生理量也是一一对应的。

#### 2.1.1 辐射量

在辐射度学中，常用的物理量简述如下：

- 辐射能（radian energy）— 辐射能  $Q_e$  指的是以电磁辐射形式发射、传播或接收的能量，单位为 J（焦耳）。当辐射能被其它物质吸收时，可以转变为其它形式的能量，如热能、电能等。
- 辐射通量（radian flux）— 辐射通量（简称辐通量） $\Phi_e$  定义了以电磁辐射形式发射、传播或接收的功率，也即单位时间内的辐射能，单位为 W（瓦）或 J/s（焦耳每秒）。
- 辐射强度（radian intensity）— 辐射强度  $I_e$  定义了辐射源在给定方向上单位立体角内的辐通量，即  $I_e = d\Phi_e / d\omega$ ，单位为 W/sr（瓦每球面度）。式中  $d\omega$  表示立体角微元。实际光源的辐射强度通常是各向异性的，可用球坐标进行表示，即  $I_e = I_e(\theta, \phi)$ 。
- 辐射出射度（radian exitance）与辐射照度（irradiance）— 辐射出射度（简称辐出度） $M_e$  与辐射照度（简称辐照度） $E_e$  均属于辐通量面密度（radian flux surface density），即辐射源（或被辐射源）表面单位面积内发出（或接收）的辐通量：

$$M_e(\text{or } E_e) = \frac{d\Phi_e}{dA} \quad (2.1)$$

式中  $dA$  表示面积微元。辐出度与辐照度的单位均为  $\text{W/m}^2$ （瓦每平方米），两者的区别在于，前者表示的是面辐射源向外发射的辐射特性，而后者表示的是辐射接收面所接收的辐

射特性。对于受照后成为面辐射源的表面而言，其辐出度显然正比于辐照度： $M_e = \rho E_e$ ，其中  $\rho$  为小于 1 的反射系数。在下一小节中还将看到，对于大部分表面而言， $\rho$  是一个波长相关的函数。

- 辐射亮度 (radiance) — 辐射亮度 (简称辐亮度)  $L_e$  定义了辐射源在给定方向上单位投影面积内的辐射强度： $L_e = dI_e / (dA \cos \theta) = d^2\Phi_e / (d\omega dA \cos \theta)$ ，单位为  $\text{W}/(\text{sr} \cdot \text{m}^2)$  (瓦每球面度每平方米)。式中  $\theta$  表示待计算位置所对应的方向矢量与辐射源表面法线之间的夹角。对于一般的辐射源而言各方向上的辐亮度通常是不同的，对于此类辐射源，其辐亮度的一般表达式为

$$L_e(\theta, \varphi) = \frac{I_e(\theta, \varphi)}{dA \cos \theta} = \frac{d^2\Phi_e(\theta, \varphi)}{d\omega dA \cos \theta} \quad (2.2)$$

将式 (2.2) 与上文中关于辐射强度  $I_e$  的定义进行联立，可得

$$dI_e(\theta, \varphi) = L_e(\theta, \varphi) dA \cos \theta \quad (2.3)$$

若某一辐射源在各方向上的辐亮度为一常数，即  $L_e(\theta, \varphi) \equiv L_e$ ，则有

$$I_e(\theta) = \frac{I_0}{dA} dA \cos \theta = I_0 \cos \theta \quad (2.4)$$

式中  $I_0$  表示面元法  $dA$  线方向上的辐射强度。此类具有各向同性辐亮度的辐射源称为余弦辐射体或朗伯体 (Lambertian source)。只有黑体才是理想的朗伯体，但一些表面粗糙的自发光体以及反射表面也可被近似视作朗伯体。

此外，根据式 (2.2) 可以计算得到面元  $dA$  在有限立体角  $\Omega$  内发出的辐通量为

$$d\Phi_e = \int d^2\Phi_e(\theta, \varphi) = \int_{\Omega} L_e(\theta, \varphi) dA \cos \theta d\omega \quad (2.5)$$

若面元为朗伯体，则对其法线指向的上半球面 ( $\Omega = 2\pi \text{ sr}$ ) 进行积分，有

$$\begin{aligned} d\Phi_e &= L_e dA \int_{2\pi} \cos \theta d\omega \\ &= \pi L_e dA \end{aligned} \quad (2.6)$$

关于式 (2.6) 的详细推导过程请参见附录一。

表 2-1 对上述辐射量的定义进行了总结，并与第 2.1.3 小节中即将介绍的光度量进行了比较。

### 2.1.2 光谱辐射量

任一实际辐射源发出的电磁辐射在不同波长上通常具有不同的能量，因此上文中介绍

的辐射量都有着与其对应的光谱辐射量，例如辐通量对应了光谱辐通量（spectral radiant flux），辐照度对应了光谱辐照度（spectral irradiance），以此类推。光谱辐射量的定义与对应的辐射量的定义非常类似，两者唯一的区别在于，辐射量是一个标量，而光谱辐射量是关于波长 $\lambda$ 的函数，其表示的是在波长 $\lambda$ 处的单位波长间隔内的辐射量。例如，光谱辐通量 $\Phi_e(\lambda)$ 表示的是辐射源发出的能量在波长 $\lambda$ 处的单位波长间隔内的辐通量。其余光谱辐射量的定义也可相应地进行推断，本小节中不再赘述。由于光谱辐射量是关于波长的函数，其表示均由相应的辐射量符号加上自变量 $\lambda$ 构成，如 $\Phi_e(\lambda)$ 、 $E_e(\lambda)$ 、 $L_e(\lambda)$ 等，其量纲也在对应辐射量的基础上增加了 $\text{nm}^{-1}$ （每纳米），如 $\text{W/nm}$ 、 $\text{W}/(\text{m}^2 \cdot \text{nm})$ 、 $\text{W}/(\text{sr} \cdot \text{m}^2 \cdot \text{nm})$ 等。

对于所有的辐射量来说，均有

$$X_e = \int_0^\infty X_e(\lambda) d\lambda \quad (2.7)$$

需要注意的是，上一小节中介绍辐出度与辐照度时所定义的反射系数 $\rho$ ，在光谱辐射量中对应了光谱反射比 $\rho(\lambda)$ ，因此有

$$M_e(\lambda) = \rho(\lambda) E_e(\lambda) \quad (2.8)$$

对于非自发光且非荧光材料的反射表面而言， $\rho(\lambda)$ 在各个波长上的数值均小于或等于 1。

对于近似于朗伯体的表面，联立式 (2.1) 与 (2.6)，并结合式 (2.8)，可得到其光谱辐亮度与光谱辐照度间的近似关系<sup>[79]</sup>：

$$L_e(\lambda) \approx \frac{\rho(\lambda)}{\pi} E_e(\lambda) \quad (2.9)$$

在研究人眼视觉系统或数字成像系统中某些不需考虑绝对数值大小的响应特性时，研究者往往并不关心光谱辐射量的具体定义方式以及量纲，而是笼统地使用“光谱功率分布函数（spectral power distribution function）”这一说法对辐射体的光谱辐射特性进行表示。根据具体应用的不同，光谱功率分布函数可以指代以上介绍的任一光谱辐射量。本论文中，我们使用符号 $P(\lambda)$ 对光谱功率分布函数进行表示。

### 2.1.3 光度量

对人眼视觉系统的研究表明，视网膜上的视觉细胞对于可见光不同波长的能量具有不同的灵敏度，其中对于黄绿光的灵敏度最高，而对于红、蓝、紫光的灵敏度则较低。人眼视觉系统的感光特性同时还会受到观察环境中亮度的影响。在明视觉 (photopic vision) 条

件下，视网膜上的杆状细胞中光化学反应达到饱和，对光不再响应，此时只有锥状细胞在工作；而在暗视觉（scotopic vision）条件下，则由杆状细胞主导光能量的感知<sup>[67]</sup>。国际照明委员会（CIE）推荐采用明视觉和暗视觉光谱光视效率函数  $V(\lambda)$  与  $V'(\lambda)$  作为理想光度观察者的光谱灵敏响应函数，并称其为标准光度观察者，如图 2-1 所示。

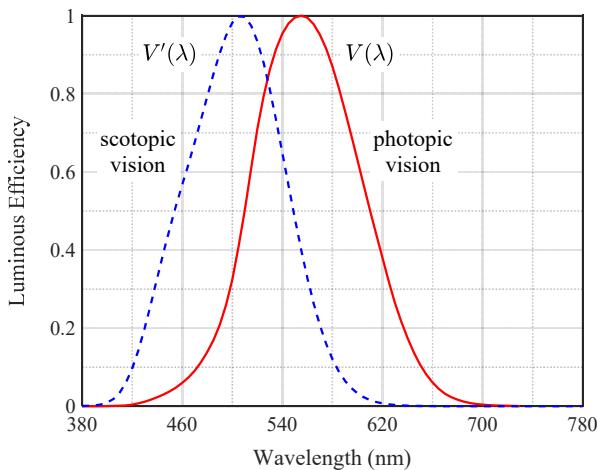


图 2-1 人眼视觉系统的明视觉与暗视觉光谱光视效率曲线

本论文中仅对明视觉条件下的颜色感知与成像机理进行探讨，故下文中与人眼视觉系统有关的光视效率均特指明视觉光谱光视效率函数  $V(\lambda)$ 。

由于光度量反映的是人眼对光辐射能的视觉强度感知，因此其与辐射量是一一对应的。两者所用的符号表示也相同，但辐射量常用下标“e”表示，而光度量则用下标“v”表示。常用的光度量包括：

- 光量（luminous energy）——光量  $Q_v$  即光的数量，其定义为光通量与照射时间的乘积，单位为  $\text{lm}\cdot\text{s}$ （流明秒）。光量的概念与辐射度学中的辐射能相当，均表示能量的多少。
- 光通量（luminous flux）——光通量  $\Phi_v$  定义了光源在单位时间内发出的光量。在光度学中，光通量是从辐通量的概念中导出的，它明确地定义为用光谱光视效率函数评价辐通量时得到的量，即，能够被人眼视觉系统所感受到的那部分辐射功率的大小。光通量的单位是  $\text{lm}$ （流明）。
- 发光强度（luminous intensity）——发光强度  $I_v$  定义了光源在指定方向上单位立体角内的光通量，单位为  $\text{cd}$ （坎德拉）。实际光源的发光强度通常是各向异性的，若用  $\theta$  表示光源空间中各方向对应的矢量与光源法线间的夹角，则  $I_v(\theta)$  在极坐标平面中构成的曲线称为辐射源的发光强度分布曲线，照明工程中也称之为配光曲线。

- 光出射度 (luminous exitance) 与光照度 (illuminance) — 光出射度  $M_v$  与光照度  $E_v$  均属于光通量面密度 (luminous flux surface density)，单位为  $\text{lm}/\text{m}^2$  (流明每平方米)。两者的区别在于，光出射度描述的是面光源向外发出的光度特性，而光照度描述的是光接收面所接收的光度特性。特别地，描述光照度时也常用  $\text{lx}$  (勒克斯) 这一单位。 $1 \text{ lx}$  定义为  $1 \text{ lm}$  的光通量均匀地照射在  $1\text{m}^2$  的面积上时的光通量面密度，即  $1 \text{ lx} = 1 \text{ lm}/\text{m}^2$ 。同样，对于受照后成为面光源的表面而言，其光出射度正比于其光照度： $M_v = \rho E_v$ 。
- 光亮度 (luminance) — 光亮度  $L_v$  定义为光源在给定方向上单位投影面积内的发光强度，单位为  $\text{cd}/\text{m}^2$  (坎德拉每平方米) 或 nit (尼特)，其中  $1 \text{ nit} = 1 \text{ cd}/\text{m}^2$ 。光亮度不仅可用来描述一个发光面，还可以用来描述光路中的任意一个截面，如一个透镜的有效面积、一个光阑所截的面积、一个像的面积，等等。

表 2-1 中列出了常见辐射量与光度量的量纲及定义。

表 2-1 基本辐射量与光度量的量纲及定义

辐射量	量纲	定义	光度量	量纲	定义
辐射能 $Q_e$	J	—	光能 $Q_v$	$\text{J}\cdot\text{s}$	—
辐通量 $\Phi_e$	W	$\Phi_e = \frac{dQ_e}{dt}$	光通量 $\Phi_v$	lm	$\Phi_v = \frac{dQ_v}{dt}$
辐射强度 $I_e$	$\text{W}/\text{sr}$	$I_e = \frac{d\Phi_e}{d\omega}$	光强度 $I_v$	cd	$I_v = \frac{d\Phi_v}{d\omega}$
辐出度 $M_e$	$\text{W}/\text{m}^2$	$M_e = \frac{d\Phi_e}{dS}$	光出射度 $M_v$	$\text{lm}/\text{m}^2$	$M_v = \frac{d\Phi_v}{dS}$
辐照度 $E_e$	$\text{W}/\text{m}^2$	$E_e = \frac{d\Phi_e}{dS}$	光照度 $E_v$	$\text{lx} = \text{lm}/\text{m}^2$	$E_v = \frac{d\Phi_v}{dS}$
辐亮度 $L_e$	$\text{W}/(\text{sr}\cdot\text{m}^2)$	$L_e = \frac{d^2\Phi_e}{d\omega dS \cos\theta}$	光亮度 $L_v$	$\text{nit} = \text{cd}/\text{m}^2$	$L_v = \frac{d^2\Phi_v}{d\omega dS \cos\theta}$

#### 2.1.4 辐射量与光度量之间的转换

由于人眼只能感受可见光波段 (380~780nm) 内的辐射能量，所以光度量就是该波段的辐射量所产生的视觉刺激值。

在明视觉条件下，辐射量转换为光度量可通过如下公式表达：

$$X_v = K_m \int_{380}^{780} X_e(\lambda) V(\lambda) d\lambda, \quad (2.10)$$

式中， $X_e(\lambda)$  表示任一光谱辐射量， $X_v$  表示与之对应的光度量。 $K_m = 683 \text{ lm/W}$  为明视觉的最大光谱光视效率，对应光视效率曲线  $V(\lambda)$  的峰值波长  $550\text{nm}$ 。

在实际计算时，往往利用求和的方法代替积分，此时式 (2.10) 可写为

$$X_v = K_m \Delta\lambda \sum_{380}^{780} X_e(\lambda) V(\lambda), \quad (2.11)$$

式中  $\Delta\lambda$  为波长间隔，通常可取为  $1\text{nm}$ 、 $5\text{nm}$  或  $10\text{nm}$ 。

## 2.2 色度学基础知识

### 2.2.1 CIE 标准色度系统

研究者们通过对人眼视觉系统进行分析发现，视网膜中存在着三种不同类型的锥体细胞，它们对不同的光谱具有不同的吸收。三种锥体细胞受到同一光刺激时产生不同的反应，这种反应的总和便形成了颜色感知。为了建立定量化和标准化的色度系统，颜色研究者们以 Grassmann 加混色定律<sup>[81]</sup>为基础进行了大量的视觉匹配实验，证明了任何颜色均可通过三个线性无关的原色按照特定比例混合而成，同时也为基础色度学提供了数据支持。

#### 2.2.1.1 CIE XYZ 标准色度系统

国际照明委员会 (CIE) 在 1931 年采用了 Wright<sup>[82]</sup>与 Guild<sup>[83]</sup>的光谱色匹配实验结果作为用来匹配等能光谱色的 RGB 光谱三刺激值，并正式推荐了 CIE1931 RGB 系统标准色度观察者色匹配函数<sup>[84, 85]</sup>。然而，在由 CIE1931 RGB 色度系统计算颜色的三刺激值时有可能出现负值，这给大量的数据处理带来了不便<sup>[67]</sup>。为此，CIE 在 CIE1931 RGB 色度系统的基础上使用假想的三原色  $X$ 、 $Y$ 、 $Z$  取代了  $700\text{nm}$ 、 $546.1\text{nm}$  和  $438.8\text{nm}$  的  $R$ 、 $G$ 、 $B$  单色光，建立并推荐了 CIE1931 XYZ 标准色度系统<sup>[86]</sup>。CIE1931 XYZ 色度系统所使用的标准色度观察者色匹配函数通常用  $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 、 $\bar{z}(\lambda)$  进行表示。

CIE1931 XYZ 标准色度系统通常适用于  $1^\circ \sim 4^\circ$  的观察视场范围。为了完善标准色度系统在大视场条件下的匹配准确性，CIE 在 1964 年根据 Stiles、Burch<sup>[87]</sup>与 Speranskaya<sup>[88]</sup>在  $10^\circ$  视场条件下获得的实验数据，推荐了 CIE1964 XYZ 补充标准色度系统。CIE1964 XYZ 色度系统所使用的标准色度观察者色匹配函数通常用  $\bar{x}_{10}(\lambda)$ 、 $\bar{y}_{10}(\lambda)$ 、 $\bar{z}_{10}(\lambda)$  进行表示。

如图 2-2 所示，两组标准色度观察者色匹配函数的形状与分布基本一致，仅在某些波段内存在差异。

以 CIE1931 XYZ 色度系统为例，对于光谱分布函数为  $\varphi(\lambda)$  的物体表面，可通过公式

$$\begin{cases} X = k \sum_{380}^{780} \varphi(\lambda) \bar{x}(\lambda) \Delta\lambda \\ Y = k \sum_{380}^{780} \varphi(\lambda) \bar{y}(\lambda) \Delta\lambda \\ Z = k \sum_{380}^{780} \varphi(\lambda) \bar{z}(\lambda) \Delta\lambda \end{cases} \quad (2.12)$$

计算其色度三刺激值。对于自发光的物体表面， $\varphi(\lambda)$  表示其光谱功率分布 (spectral power distribution) 函数，即  $\varphi(\lambda) = P(\lambda)$ ；对于受照后二次发光的物体表面， $\varphi(\lambda)$  则表示其光谱反射比（或透射比）函数与光源光谱功率分布函数的乘积<sup>[84]</sup>，即  $\varphi(\lambda) = \rho(\lambda)P(\lambda)$  或  $\varphi(\lambda) = \tau(\lambda)P(\lambda)$ ，其中  $\rho(\lambda)$  与  $\tau(\lambda)$  分别代表物体表面的光谱反射比与光谱透射比函数。

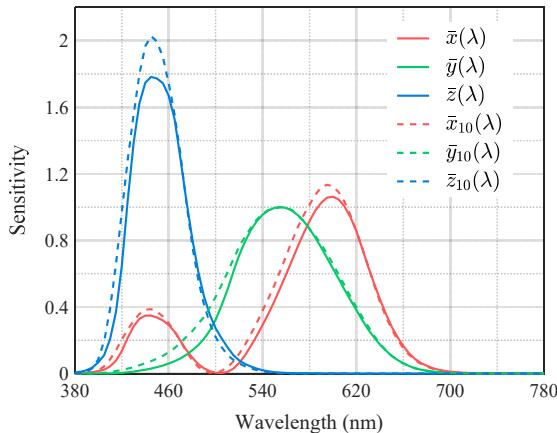


图 2-2 CIE1931 XYZ (2°) 与 CIE1964 XYZ (10°) 标准色度观察者色匹配函数的比较

式 (2.12) 中的  $k$  代表 CIE1931 XYZ 色度系统所对应的归一化系数。当计算绝对三刺激值（量纲与光亮度相同，均为  $\text{cd}/\text{m}^2$ ）时， $k = 683 \text{ lm}/\text{W}$ ，此时  $\varphi(\lambda)$  应为绝对辐亮度，单位为  $\text{W}/(\text{sr} \cdot \text{m}^2 \cdot \text{nm})$ ；当计算相对三刺激（无量纲）时，可通过令光源的  $Y$  值等于 100 以实现归一化，即

$$k = \frac{100}{\sum_{380}^{780} \varphi(\lambda) \bar{y}(\lambda) \Delta\lambda} \quad (2.13)$$

在某些与图像处理有关的应用中，出于计算方便的考虑，有时也将光源的  $Y$  值归一化为 1，此时只需将式 (2.13) 中的分子替换为 1 即可。本论文中，若无特殊说明，则  $X$ 、 $Y$ 、 $Z$  值均指归一化至  $[0, 1]$  区间内的相对三刺激值。

为方便计算，式 (2.12) 常以矩阵乘法的形式进行表达：

$$\mathbf{X} = k\Phi \cdot \mathbf{S}_{\text{cmf}}^T \quad (2.14)$$

式中  $\mathbf{X}=[X, Y, Z]$  表示三刺激值向量， $\Phi=[\varphi(380), \dots, \varphi(780)]$  表示光谱分布函数对应的  $1 \times N$  行向量 ( $N$  为波长采样数)， $\mathbf{S}_{\text{cmf}}^T$  表示  $N \times 3$  标准观察者色匹配函数矩阵，其每一列对应一组色匹配函数，即

$$\mathbf{S}_{\text{cmf}}^T = \begin{bmatrix} \bar{x}(380) & \bar{y}(380) & \bar{z}(380) \\ \vdots & \vdots & \vdots \\ \bar{x}(780) & \bar{y}(780) & \bar{z}(780) \end{bmatrix} \quad (2.15)$$

### 2.2.1.2 CIE xy 色品坐标平面

CIEXYZ 色度系统以三维向量的形式表示颜色，而在某些场合中，人们更关心颜色的色品 (chromaticity) 而非其绝对强度。色品坐标允许我们在二维平面内对 XYZ 三刺激值之间的相对关系进行描述：

$$\begin{cases} x = \frac{X}{X+Y+Z} \\ y = \frac{Y}{X+Y+Z} \\ z = \frac{Z}{X+Y+Z} = 1-x-y \end{cases} \quad (2.16)$$

式中由  $x$ 、 $y$  构成的平面通常被称为 CIE1931 (或 CIE1964) 色品坐标平面。通过将可见光波段内所有单色光在  $(x, y)$  平面上对应的坐标连接起来，即可得到 CIE1931 (或 CIE1964) 色度系统下的二维色域边界，如图 2-3 所示。

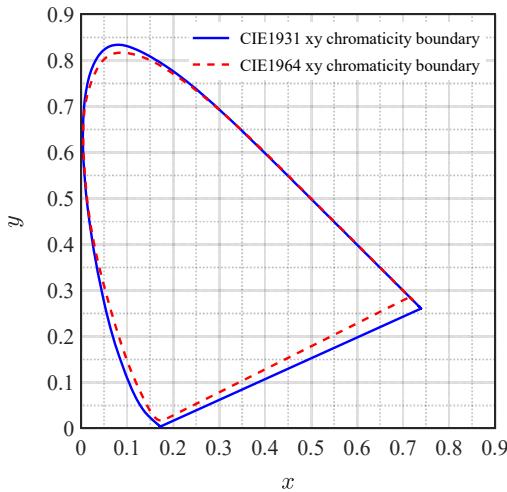


图 2-3 CIE1931 与 CIE1964 色度系统所对应的二维色域边界

CIE XYZ 标准色度系统是国际上对颜色进行定量描述的统一标准，也是几乎所有颜色测量与色度计算的理论基础。

## 2.2.2 均匀颜色空间与色差评价

### 2.2.2.1 均匀颜色空间

CIE1931 与 CIE1964 标准色度系统的提出允许人们方便地对颜色进行描述，然而该系统存在明显的感知非均匀性，换句话说，XYZ 颜色空间中欧氏距离的大小与人眼视觉系统实际感知色差的大小并不吻合。为解决这一问题，CIE 在大量视觉实验的基础上于 1960 年与 1976 年相继提出了两个版本的均匀色品标尺图（uniform-chromaticity-scale diagram, UCS），简称为 CIE1960 UCS 以及 CIE1976 UCS<sup>[67]</sup>。

然而，UCS 中的色品坐标仅仅是 XYZ 色度系统中  $x$ 、 $y$  色品坐标的扩展，只能反映相对色度量，而无法完整地描述颜色空间中所有的感知属性。为此，CIE 在整理和总结了当时出现的诸多颜色空间的基础上，于 1976 年正式推荐了 CIE1976  $L^* a^* b^*$ （简称 CIELAB）颜色空间。该颜色空间显著改善了视觉非均匀性，可以更为准确地对色差进行定量评价。

CIELAB 颜色空间采用明度  $L^*$  和色品坐标  $a^*$ 、 $b^*$  作为三维正交分量：

$$\begin{cases} L^* = 116f(Y/Y_n) - 16 \\ a^* = 500[f(X/X_n) - f(Y/Y_n)] \\ b^* = 200[f(Y/Y_n) - f(Z/Z_n)] \end{cases} \quad (2.17)$$

式中  $X$ 、 $Y$ 、 $Z$  表示颜色样品的三刺激值， $X_n$ 、 $Y_n$ 、 $Z_n$  表示参考照明体经完全漫反射体（光谱反射比函数在各波长处恒为 1 的理想反射表面）反射后的三刺激值，且有

$$f(t) = \begin{cases} t^{1/3} & t > 0.008856 \\ 7.787t + 16/116 & t \leq 0.008856 \end{cases} \quad (2.18)$$

若将  $L^*$ 、 $a^*$ 、 $b^*$  在柱坐标系中进行表示，则可以得到与心理量近似对应的一组颜色感知属性，即明度（lightness） $L^*$ 、彩度（chroma） $C_{ab}^*$  以及色调角（hue angle） $h_{ab}$ 。以  $a^*b^*$  色品平面第一象限为例，彩度与色调角的计算公式如下：

$$\begin{cases} C_{ab}^* = \sqrt{(a^*)^2 + (b^*)^2} \\ h_{ab} = \tan^{-1}(b^*/a^*) \end{cases} \quad (2.19)$$

### 2.2.2.2 色差评价

用来定量地表示颜色感知差别的量值称为色差（color difference）。基于大量实验获取的颜色视觉判断及色差比较心理物理数据可以帮助建立用于预测视觉感知差异的色差公式。目前，国际上较为通用的色差计算方法包括 CIELAB 色差公式以及 CIEDE2000 色差公式。

CIELAB 色差公式直接使用一对颜色刺激在 CIELAB 颜色空间中的欧氏距离作为色差数值大小：

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (2.20)$$

式中  $\Delta L^*$ 、 $\Delta a^*$ 、 $\Delta b^*$  分别表示两个颜色刺激在 CIELAB 空间内的  $L^*$ 、 $a^*$ 、 $b^*$  坐标之差。

CIE 在 2001 年正式推荐了 CIEDE2000 色差公式，其在 CIELAB 色差公式的基础上进一步改善了色差计算的空间均匀性<sup>[89]</sup>。CIEDE2000 色差公式的具体计算过程如下<sup>[90]</sup>：

- 1) 依照式 (2.17) ~ (2.19)，分别计算待比较的一对色样的  $L^*$ 、 $a^*$ 、 $b^*$ 、 $C_{ab}^*$  值；
- 2) 分别计算这对色样的  $a'$ 、 $C'$ 、 $h'$  值：

$$\begin{cases} L' = L^* \\ a' = (1+G)a^* \\ b' = b^* \\ C' = \sqrt{(a')^2 + (b')^2} \\ h' = \tan^{-1}(b'/a') \end{cases} \quad (2.21)$$

其中

$$G = 0.5 \left( 1 - \sqrt{\frac{(\bar{C}^*)^7}{(\bar{C}^*)^7 + 25^7}} \right) \quad (2.22)$$

式中  $\bar{C}^*$  表示两个色样  $C_{ab}^*$  值的算数平均值；

3) 计算  $\Delta L'$ 、 $\Delta C'$ 、 $\Delta H'$  值:

$$\begin{cases} \Delta L' = L'_b - L'_s \\ \Delta C' = C'_b - C'_s \\ \Delta H' = 2\sqrt{C'_b C'_s} \sin[(h'_b - h'_s)/2] \end{cases} \quad (2.23)$$

式中下标“*b*”表示色样对中的试样，“*s*”表示标样；

4) 计算 CIEDE2000 色差  $\Delta E_{00}$ :

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{k_C S_C}\right) \left(\frac{\Delta H'}{k_H S_H}\right)} \quad (2.24)$$

式中

$$\begin{cases} S_L = 1 + \frac{0.015(\bar{L}' - 50)^2}{[20 + (\bar{L}' - 50)^2]^{1/2}} \\ S_C = 1 + 0.045\bar{C}' \\ S_H = 1 + 0.015\bar{C}'T \\ R_C = 2 \left( \frac{\bar{C}'^7}{\bar{C}'^7 + 25^7} \right)^{1/2} \\ T = 1 - 0.17 \cos(\bar{h}' - 30^\circ) + 0.24 \cos(2\bar{h}') + 0.32 \cos(3\bar{h}' + 6^\circ) - 0.2 \cos(4\bar{h}' - 63^\circ) \\ R_T = -\sin(2\Delta\theta)R_C \\ \Delta\theta = 30 \exp\left[-\frac{(\bar{h}' - 275^\circ)^2}{25^2}\right] \end{cases} \quad (2.25)$$

$\bar{L}'$ 、 $\bar{C}'$ 、 $\bar{h}'$  分别为两个色样  $L'$ 、 $C'$ 、 $h'$  的算数平均值， $k_L$ 、 $k_C$ 、 $k_H$  为权重因子，用于调整分色差的权重从而对不同观察条件进行修正，通常情况下均设为 1。

CIEDE2000 色差公式的计算方法较为复杂，但相比 CIELAB 色差公式具有更高的预测精度，弥补了后者在预测小色差时与视觉感知差异之间的不一致性。不过，CIELAB 色差公式因计算简单方便，至今仍被广泛使用。

### 2.2.3 光源与色温

#### 2.2.3.1 光源与标准照明体

光源是使人眼视觉系统产生视觉感知的必要条件，也是影响物体颜色外貌的重要因素之一。颜色科学中通常使用光谱功率分布函数  $P(\lambda)$  对光源进行定量描述。为了统一颜色评价的标准，便于比较和传递，CIE 针对颜色的测量和计算推荐了几种标准照明体和标准光源，包括标准照明体 A、B、C、D、E 以及标准光源 A、B、C。需要说明的是，根据 CIE

的定义，“光源”是指能发光的物理辐射体，如灯、太阳等；而“照明体”仅需具有特定的光谱功率分布函数，而这种分布函数不是必须由一个具体的光源直接提供或实现，它可以由表格的形式给出<sup>[67]</sup>。

对于图像处理应用而言，在 CIE 推荐的照明体与光源中，较常使用的有标准照明体 A (2856K 黑体)、D50 (5003K 重组日光)、D55 (5503K 重组日光)、D65 (6504K 重组日光)、D75 (7504K 重组日光)、E (等能白光)、CWF (4230K 冷白荧光灯)、F8 (5000K 荧光灯日光模拟器)、TL84 (4000K 欧洲商店模拟光源) 等，其相对光谱功率分布函数如图 2-4 所示。

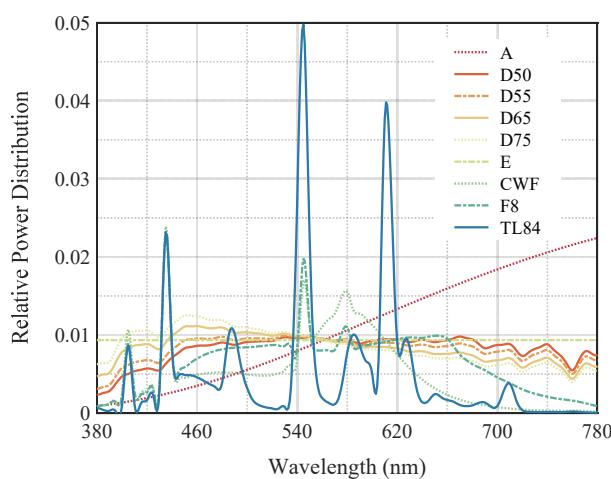


图 2-4 九种 CIE 标准照明体的相对光谱功率分布函数（均归一化至  $Y=1$ ）

### 2.2.3.2 色温与相关色温

根据 Planck 公式，黑体辐射的光谱功率分布函数完全由其温度决定<sup>[91]</sup>。通过将不同温度黑体的色品坐标点在色品平面上连接起来便形成了黑体轨迹线 (Planckian locus)，如图 2-5 中黑色实线所示。对于某些受热辐射的发光体 (如白炽灯)，当其在温度  $T$  时所呈现的颜色与黑体在温度  $T_c$  时的颜色相同时，可将黑体的温度  $T_c$  称为该辐射源的颜色温度，简称色温 (color temperature)。色温的单位与绝对温度的单位相同，均为 K (开尔文)。

对于发光机理并非受热辐射的其它光源 (如荧光灯、LED、气体放电光源等)，其光谱功率分布有可能与黑体大相径庭，所对应的色品坐标也不一定准确地落在黑体温度轨迹上。对于此类光源，不再适合使用色温进行描述，而应该采用相关色温的概念来进行光色的表征和比较。相关色温 (correlated color temperature, CCT) 被定义为黑体轨迹上与待测光源色品最为接近的坐标点所对应的色温<sup>[92]</sup>。在上文中提及的 9 种标准照明体中，除 A 照明

体之外，其余照明体所标示的温度均为相关色温，图 2-5 展示了它们的色品坐标与黑体轨迹的相对位置关系。

对于色品坐标与黑体色温轨迹相距不太远的光源，通常可以认为，色温越低，其呈现出的颜色外貌越偏向红色、橙色等暖色调；而随着色温的升高，其颜色外貌逐渐向蓝色、青色等冷色调转变。

人眼视觉系统中的颜色恒常性 (color constancy) 使得人们能够很快地对所处环境中光源的颜色产生适应。然而，数码相机并不具备这种光源适应能力，其对于物体颜色的感知完全由物体反射的光谱特性所决定。如第 2.2.1 小节所述，物体反射出的光谱能量同时取决于光源的光谱功率分布函数以及物体自身的光谱反射比函数，因此，当光源色温发生变化时，数码相机所捕获的原始图像也会发生明显的色调变化。图 2-6 展示了同一场景在不同色温光源下的颜色外貌变化（并非真实的原始图像，仅用作示意）。

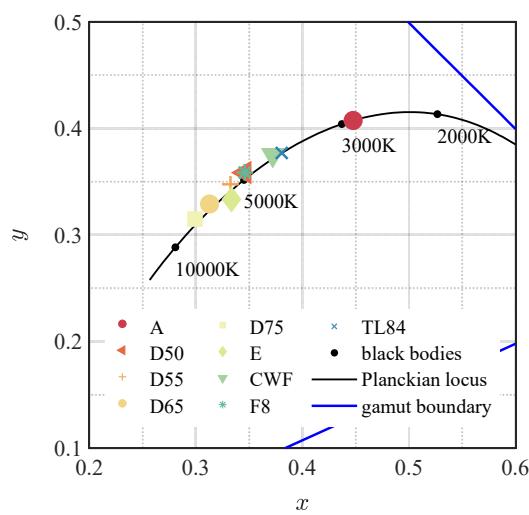


图 2-5 黑体色温轨迹及 9 种 CIE 标准照明体的色品平面坐标



图 2-6 数码相机对于同一场景在低（左）、中、高（右）色温光源照明下捕获的图像（未经白平衡校正）

## 2.3 色貌现象与色适应模型

CIE 标准色度系统的建立主要基于特定照明环境和观察条件下的色块视匹配实验，因此一对具有相同 XYZ 三刺激值的物体表面，只有在照明条件、观察条件、尺寸、形状、表面特性甚至背景环境等因素完全相同的情况下，才能对人眼视觉系统产生相同的视觉感知。一旦这些限制条件无法得到满足时，即使相同的两组三刺激值也有可能令观察者产生不同的视觉感知，这种现象通常被称为色貌（color appearance）现象<sup>[93]</sup>。

为了对不同媒介在不同的观察条件与观察环境下的真实颜色感知现象进行描述、解释以及预测，CIELAB 模型<sup>\*</sup>、Nayatani 模型<sup>[94]</sup>、Hunt 模型<sup>[95]</sup>、RLAB 模型<sup>[96]</sup>等各种色貌模型相继被提出。2004 年，CIE 在 CIECAM97s 色貌模型的基础上提出了更为简洁且高效的 CIECAM02 色貌模型，并将其作为工业应用中的推荐色貌模型<sup>[97]</sup>。

本节中我们仅对 CIECAM02 色貌模型中的核心部分——CAT02 色适应模型——进行介绍，关于 CIECAM02 模型的更多细节请参阅<sup>[97]</sup>与<sup>[93]</sup>。

色适应（chromatic adaptation）指的是人眼视觉系统在照明条件发生变化时能够自适应地对视网膜上锥体细胞的响应灵敏度进行调节从而形成稳定的物体颜色感知的现象。色适应是色貌现象中颜色恒常性产生的根本原因，也是建立色貌模型的核心基础。通过对色适应过程进行定量描述，我们可建立起用于连接不同媒介或不同照明观察条件下的颜色感知的变换模型。

CIECAM02 色貌模型使用色适应变换的推荐模型 CAT02 对不同光源、不同观察条件下人眼视觉系统的颜色感知进行预测。

CAT02 色适应变换模型以待适应光源下的物体色三刺激值  $[X, Y, Z]$ 、待适应光源三刺激值  $[X_w, Y_w, Z_w]$ 、参考光源三刺激值  $[X_w^{\text{ref}}, Y_w^{\text{ref}}, Z_w^{\text{ref}}]$ 、适应场亮度  $L_A$  以及背景相对亮度等级作为输入，以色适应后参考光源下对应色（corresponding color，指的是在不同照明条件下具有相同感知颜色外貌的一对颜色刺激）的三刺激值  $[X_c, Y_c, Z_c]$  作为输出。在给定背景相对亮度等级（平均（average）、昏暗（dim）以及黑暗（dark））后，CAT02 模型首先对环境影响非线性参数（exponential nonlinearity） $c$ 、色诱导因子（chromatic induction factor） $N_c$  以及最大适应程度（maximum degree of adaptation） $F$  进行计算，如表 2-2 所示。

---

<sup>\*</sup> CIELAB 颜色空间由于具有明度、彩度和色调角等感知颜色属性量，因此也被认为是一个简单的色貌模型。

表 2-2 CAT02 色适应变换模型中各参数与背景相对亮度等级的关系

背景相对亮度等级	非线性参数 $c$	色诱导因子 $N_c$	最大适应程度 $F$
平均 (average)	0.69	1.0	1.0
昏暗 (dim)	0.59	0.9	0.9
黑暗 (dark)	0.525	0.8	0.8

利用适应场亮度  $L_A$  与最大适应程度  $F$ , 可对色适应程度  $D$  进行计算:

$$D = F \left[ 1 - \left( \frac{1}{3.6} \right) e^{-\frac{L_A + 42}{92}} \right] \quad (2.26)$$

对于输入的各项 XYZ 三刺激值 (需归一化至 [0,100] 范围), CAT02 模型使用变换矩阵  $\mathbf{M}_{\text{CAT02}}$  将其转换为一组特定颜色空间中的 RGB 值:

$$\begin{cases} [R, G, B]^T = \mathbf{M}_{\text{CAT02}} \cdot [X, Y, Z]^T \\ [R_w, G_w, B_w]^T = \mathbf{M}_{\text{CAT02}} \cdot [X_w, Y_w, Z_w]^T \\ [R_w^{\text{ref}}, G_w^{\text{ref}}, B_w^{\text{ref}}]^T = \mathbf{M}_{\text{CAT02}} \cdot [X_w^{\text{ref}}, Y_w^{\text{ref}}, Z_w^{\text{ref}}]^T \end{cases} \quad (2.27)$$

式中

$$\mathbf{M}_{\text{CAT02}} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \quad (2.28)$$

接着, 在特定的 RGB 颜色空间中对色适应后参考光源下的对应色刺激值  $[R_c, G_c, B_c]$  进行计算:

$$\begin{cases} R_c = \left[ \frac{Y_w R_w^{\text{ref}}}{Y_w^{\text{ref}} R_w} D + (1-D) \right] R \\ G_c = \left[ \frac{Y_w G_w^{\text{ref}}}{Y_w^{\text{ref}} G_w} D + (1-D) \right] G \\ B_c = \left[ \frac{Y_w B_w^{\text{ref}}}{Y_w^{\text{ref}} B_w} D + (1-D) \right] B \end{cases} \quad (2.29)$$

最后, 将适应后的 RGB 刺激值重新转换回 XYZ 颜色空间中:

$$[X_c, Y_c, Z_c]^T = \mathbf{M}_{\text{CAT02}}^{-1} \cdot [R_c, G_c, B_c]^T \quad (2.30)$$

此时的  $[X_c, Y_c, Z_c]$  即表示参考光源下与适应后的物体表面具有相同颜色感知的对应色的三刺激值。

概括地说, CAT02 色适应变换模型可用一个多输入单输出的函数进行表示:

$$\mathbf{X}_c = f_{\text{CAT02}}(\mathbf{X}, \mathbf{X}_w, \mathbf{X}_w^{\text{ref}}, L_A, B) \quad (2.31)$$

式中  $B \in \{\text{average}, \text{dim}, \text{dark}\}$  表示背景相对亮度等级, 粗体变量表示各自对应的 XYZ 三刺激值向量。

## 2.4 本章小结

本章首先对辐射度学与光度学的相关基础知识进行了介绍。辐射度学是一门研究光辐射能量计量的学科, 其研究对象是真实世界中客观存在的各种物理辐射量; 而光度学是一门研究人眼对辐射能量感知规律的学科, 其研究对象是使用光谱光视效率函数对不同波长的辐射能量进行加权后得到的主观心理物理量。同时, 本章也对 CIE 标准色度系统、均匀颜色空间、颜色差异的定量评价指标以及一些简单的照明工程基础知识进行了介绍, 这部分知识也是支撑本论文中大部分研究内容的理论基础。当将图像颜色外貌与人眼视觉感知之间的一致性作为图像复现结果的评价指标时, 往往还需要考虑照明条件、观察条件、背景环境等光度色度参数之外的信息, 因此本章的最后还对用于描述与预测各种视觉现象的色貌模型进行了概述。

### 第三章 数码相机原始响应值预测模型

从用户按下快门到图像在显示屏上进行展示，在这短短的数百毫秒内，数码相机完成了一次场景中的光信号至最终输出数字信号的完整转换过程。根据数据分析、处理所针对目标的不同，这一转换过程可分为两个阶段：第一阶段与摄影光度学（photographic photometry）的研究目标相对应，旨在使用数字量对客观辐射量进行忠实记录，必要时可对光信号与原始数字信号间的映射关系进行建模，从而允许用户通过图像信息获取场景中的辐射度、光度参数；第二阶段与计算摄影学（computational photography）的研究目标相对应，即通过硬件设计与软件计算的有机结合，增强或扩展数码相机的数据分析能力，从而获得符合甚至优于人眼视觉感知的输出图像。该二阶段的信号传输、处理流程如图 3-1 所示。

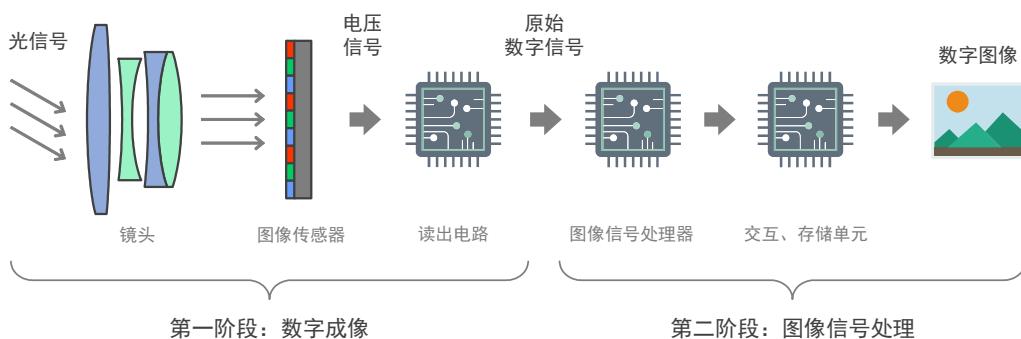


图 3-1 典型的数字成像系统信号传输流程

本章将以第一阶段中光信号至原始数字信号的映射关系作为研究目标，在对数字成像物理模型进行严格论证与分析的基础上，提出一种启发式的数码相机原始响应值预测模型，并以此为第二阶段中各种图像处理算法生成所需的标定数据。

#### 3.1 数码相机原始响应值构成模型

某一物体表面（或光源）向空间中发射的部分辐射能量被图像传感器接收进而转化为原始数字响应值，这一过程可进一步地分解为三个子过程：1) 辐射能量经光学系统传递至图像传感器感光单元表面，2) 感光元件通过光电效应将辐射能量转化为电荷信号，3) 电荷信号被转换为电压信号，并经模数转换单元转换为最终的输出数字信号。

### 3.1.1 感光单元表面光谱辐照度构成模型

首先, 以图 3-2 所示单透镜理想光学系统为例对传感器平面所接收的辐射量进行分析。当光学系统对所拍摄的物体表面准确对焦时, 物体表面微元  $dA_o$  发射的部分辐通量通过光阑进入相机内部, 其光路经透镜折射后再次汇聚于传感器平面上某一微元  $dA_s$  处。不考虑衍射等因素, 对于  $dA_s$  而言, 其接收到的所有辐射能量均来自于  $dA_o$ 。

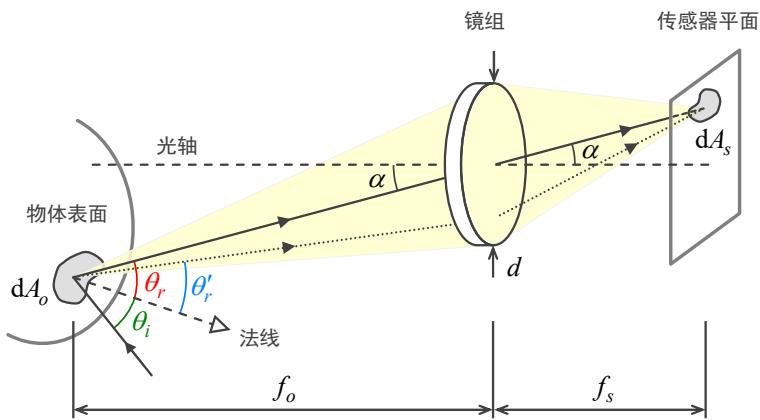


图 3-2 单透镜理想光学系统成像模型

设面元  $dA_o$  在空间中各方向上的光谱辐亮度为  $L_o(\theta, \varphi, \lambda)$ , 根据式 (2.2) 可知, 由  $dA_o$  发出并经光阑进入相机内部的光谱辐通量为

$$d\Phi(\lambda) = dA_o \int_{\Omega_o} L_o(\theta, \varphi, \lambda) \cos \theta'_r d\omega \quad (3.1)$$

式中  $\Omega_o$  表示光阑对  $dA_o$  所张的圆锥立体角 (图 3-2 中左侧浅黄色阴影部分),  $\theta'_r$  为光阑内某一点与  $dA_o$  的连线相对于  $dA_o$  法线的夹角。

若光学系统无能量衰减, 则式 (3.1) 中的辐通量将完全由  $dA_s$  接收, 因此  $dA_s$  处的光谱辐照度为<sup>[98]</sup>

$$E_s(\lambda) = \frac{d\Phi(\lambda)}{dA_s} t(\lambda) \quad (3.2)$$

式中  $t(\lambda)$  表示系统的光谱透过率函数。

根据单透镜成像模型,  $dA_o$  对光阑中心所张的立体角应与  $dA_s$  对光阑中心所张的立体角相等 (图 3-2 中未画出), 因此有

$$\frac{dA_o \cos \theta_r}{(f_o / \cos \alpha)^2} = \frac{dA_s \cos \alpha}{(f_s / \cos \alpha)^2} \quad (3.3)$$

式中  $\theta_r$  表示  $dA_o$  发出的主光线与其法线的夹角， $\alpha$  表示该主光线与相机系统光轴的夹角， $f_o$  与  $f_s$  分别表示物距与像距。

联立式 (3.1) ~ (3.3)，有

$$E_s(\lambda) = \cos \alpha (f_o^2 / f_s^2) t(\lambda) \int_{\Omega_o} L_o(\theta, \varphi, \lambda) \frac{\cos \theta'_r}{\cos \theta_r} d\omega \quad (3.4)$$

该式即为单透镜光学系统感光单元表面光谱辐照度的构成模型。

若物体与相机的距离  $f_0$  相对于光阑的尺寸  $d$  足够大，则  $\Omega_o$  为一小立体角，此时可认为光谱辐亮度在这一小立体角内与方向无关（若物体表面为朗伯体，则无需这一假设条件该命题也成立），且有  $\theta'_r \approx \theta_r$ 。在该假设下，式 (3.4) 可简化为

$$\begin{aligned} E_s(\lambda) &\approx \cos \alpha (f_o^2 / f_s^2) L_o(\lambda) t(\lambda) \Omega_o \\ &\approx \cos \alpha (f_o^2 / f_s^2) L_o(\lambda) t(\lambda) \frac{(\pi d^2 / 4) \cos \alpha}{(f_o / \cos \alpha)^2} \\ &= \frac{\pi}{4F^2} \cos^4 \alpha L_o(\lambda) t(\lambda) \end{aligned} \quad (3.5)$$

式中  $F = f_s/d$  表示光学系统的相对孔径（也称为光圈数）。

以上光谱辐照度的推导过程中仅仅考虑了理想单透镜光学系统下的情形<sup>[99]</sup>。对于实际的相机系统而言，其传感器各位置处接收的光谱辐照度除了取决于主光线与光轴的夹角  $\alpha$  之外，还会受到诸多因素的影响<sup>[100-103]</sup>。关于该部分内容的具体分析将在第 4.2.1 小节中展开，本节中仅笼统地使用  $U(i, j)$  对传感器上  $(i, j)$  坐标处的辐照度调制因子进行表示。换言之， $U(i, j)$  中既包括了式 (3.5) 中  $\cos^4 \alpha$  的径向照度衰减<sup>\*</sup>，也包括了由渐晕<sup>[104]</sup>、微透镜（micro-lens）修正<sup>[105]</sup>等因素导致的照度调制。基于此简化表示，式 (3.5) 可改写为

$$E_s(i, j, \lambda) = \frac{\pi}{4F^2} U(i, j) L(\lambda) t(\lambda) \quad (3.6)$$

式中光谱辐亮度  $L_o(\lambda)$  的下标 “o” 已被省略。

$U(i, j)$  可被视为相机系统的一种内在特性，其分布取决于光学系统的结构以及光圈、对焦距离等拍摄参数，同时还会受到拍摄内容的影响（光学系统的色差与入射光的光谱分布有关）。在实际应用中，通常采用在可控环境下拍摄参考场景的方法对  $U(i, j)$  进行预先标定。

\* 根据图 3-2 中的几何关系，易知有  $\alpha = \tan^{-1} \left[ \frac{\sqrt{(i-i_0)^2 + (j-j_0)^2}}{f_s} \right]$ ，其中  $(i_0, j_0)$  表示位于传感器中央的感光单元的坐标。

此外，需要说明的是，式 (3.4)~(3.6) 中的光谱透过率函数  $t(\lambda)$  既包括了光学模组（镜头、低通滤波片、微透镜等）的光谱透过率  $t_{\text{lens}}(\lambda)$ ，也包括了感光单元上方彩色滤色片的透过率  $t_{\text{filter}}(\lambda)$ 。对于常见的 CFA 传感器（以 Bayer 阵列<sup>[106]</sup>、X-Trans 阵列<sup>[107]</sup>等方式进行彩色滤色片排布的传感器）， $t_{\text{filter}}(\lambda)$  根据该感光单元所处的位置不同而对应于  $t_{\text{filter}}^r(\lambda)$ 、 $t_{\text{filter}}^g(\lambda)$ 、 $t_{\text{filter}}^b(\lambda)$  三种透过率函数之一；对于 Foveon 传感器<sup>[108]</sup>， $t_{\text{filter}}(\lambda)$  表示位于目标通道之上的所有透过层的组合光谱透过率。

### 3.1.2 电荷信号构成模型

暂且忽略传感器平面上的位置坐标，仅以某一感光单元（像素）作为研究对象。一次曝光周期内，一个感光单元激发并释放出的电子个数  $I$  将取决于其光敏面接收的光谱辐射能量  $Q(\lambda)$  以及光电转换器件的转换效率  $q(\lambda)$ <sup>[109]</sup>：

$$I = \int_{\lambda} Q(\lambda) q(\lambda) d\lambda \quad (3.7)$$

式中  $q(\lambda)$  的物理意义为该光电二极管在波长  $\lambda$  处每吸收 1 焦耳能量后能够释放出的电子个数，单位为  $e^-/(J \cdot nm)$ 。本节及后文中，在不引起歧义的情况下，光谱辐射量的下标“ $e$ ”将被省略。

若假设在一次曝光周期内光源与物体表面保持足够稳定，各参数均不随时间变化而变化，则有

$$Q(\lambda) = T \cdot \Phi(\lambda) \quad (3.8)$$

式中  $T$  为曝光时间， $\Phi(\lambda)$  为光敏面接收的光谱辐通量。同时，考虑到光敏面各位置处的光谱辐通量分布可能存在差异，我们使用光谱辐照度  $E(x, y, \lambda)$  与面积微元的积分对光谱辐通量进行表示，即

$$\Phi(\lambda) = \int_x \int_y E(x, y, \lambda) dx dy \quad (3.9)$$

联立式 (3.7) ~ (3.9)，有

$$I = T \int_{\lambda} \int_x \int_y E(x, y, \lambda) S(x, y) q(\lambda) dx dy d\lambda \quad (3.10)$$

式中  $S(x, y)$  表示感光单元内部的灵敏度空间波动，通常由感光介质的不均匀性引起，对于质量合格的传感器而言，可认为  $|S(x, y) - 1| \ll 1$ 。

从量纲的角度分析式 (3.10) 有助于更加直观地理解该过程：积分时间  $T$  以 s (秒) 为单位，光谱辐照度  $E(x, y, \lambda)$  以  $W/(m^2 \cdot nm)$  为单位， $S_r(x, y)$  无量纲。因此， $I$  的量纲为

$s \cdot (W \cdot m^{-2}) \cdot (e^- \cdot J^{-1}) \cdot nm^{-1} \cdot m^2 \cdot nm = e^-$ , 即释放出的电子个数。

当感光单元的尺寸足够小时, 可近似地认为其光敏面各处的光谱辐照度与位置坐标无关<sup>\*</sup>。因此, 式 (3.10) 可改写为

$$I = T \bar{S} A \int_{\lambda} E(\lambda) q(\lambda) d\lambda \quad (3.11)$$

式中  $\bar{S}$  表示  $S(x, y)$  在该感光单元内的空间平均值,  $A$  表示该感光单元的有效感光面积。

联立 (3.6) 与 (3.11), 有

$$I = \frac{\pi}{4F^2} T \bar{S} A U \int_{\lambda} L(\lambda) t(\lambda) q(\lambda) d\lambda \quad (3.12)$$

式中  $U$  表示该感光单元所对应的辐照度调制因子。

式 (3.12) 即为数码相机系统将空间光信号转化为电荷信号的完整模型。

### 3.1.3 数字信号构成模型

从感光单元释放出的电荷形成电流, 进而在转换单元两端形成电压差。该电压信号将经由放大单元进行放大:

$$V = f(g'I + V_{\text{offset}}) \quad (3.13)$$

式中  $g'$  表示信号放大单元的增益倍数 (gain), 由电路的基础增益以及拍摄时所用的感光度参数 (ISO) 共同决定;  $V_{\text{offset}}$  表示在信号放大单元输出的电压信号上人为加入的一个偏置电压, 相机设计者通过这种方式来避免小于零的输出信号被模数转换单元截零从而改变噪声的分布<sup>†</sup>;  $f(\cdot)$  表示整个模拟电路中引入的非线性变换<sup>[110]</sup>, 在信号未饱和的情况下, 除非传感器设计者有意对输出电压信号进行非线性调制(例如为了获得更高的动态范围<sup>[111, 112]</sup>), 否则对于大部分传感器来说, 可认为  $f(x) \approx x$ 。

最后, 该电压信号  $V$  进入数模转换单元 (ADC) 被转换为原始数字响应值  $D$ :

$$D = \text{round}(V/\eta) \quad (3.14)$$

式中  $\text{round}(\cdot)$  表示取整操作,  $\eta$  表示量化步长, 且通常有  $\eta = \frac{f(gI_{\max} + V_{\text{offset}})}{2^b - 1}$ ,  $I_{\max}$  为该感光器件所对应的饱和电子数(满阱容量),  $b$  为位深。原始数字响应值通常用 DN(digital number)

<sup>\*</sup> 以 4000 万像素的 36mm×24mm 传感器为例, 当搭配 50mm 焦距镜头时, 该传感器上每一个感光单元对应约 0.0026° 视场角。当物方平面与相机相距 5m 时, 该视场角对应约 0.23mm 的物体尺寸。这一尺寸小于现实场景中大部分物体表面纹理的空间变化周期, 因此在拍摄近处的物体时, 可近似认为各个感光单元内部的光谱辐照度与位置坐标无关。

<sup>†</sup> 这种人为引入的偏置量与量化步长之比  $V_{\text{offset}}/\eta$  通常也被称作暗电平 (black level) (假设  $f(x) = x$ )。

进行表示<sup>[113]</sup>, 因此  $\eta$  的量纲为 volts/DN (伏特每 DN) \*。

联立式 (3.12) ~ (3.14), 有

$$D = \text{round} \left[ f \left( \frac{\pi}{4F^2} g T \bar{S} A U \int_{\lambda} L(\lambda) t(\lambda) q(\lambda) d\lambda + D_{\text{offset}} \right) \right] \quad (3.15)$$

式中  $g = f^{-1}[f(g')/\eta]$  表示模数电路中的综合增益<sup>[114]</sup>,  $D_{\text{offset}} = f^{-1}[f(V_{\text{offset}})/\eta]$  表示响应值偏置量。

式 (3.15) 即为本论文中所采用的数码相机的原始响应值构成模型。为方便起见, 表 3-1 中列出了式 (3.15) 中各符号的含义及量纲。

表 3-1 式 (3.15) 中各符号的含义及量纲

符号	含义	量纲
$L(\lambda)$	被摄物体表面的光谱辐亮度	$\text{W}/(\text{sr} \cdot \text{m}^2 \cdot \text{nm})$
$t(\lambda)$	相机系统的综合光谱透过率函数	$\text{nm}^{-1}$
$q(\lambda)$	光电转换效率函数	$\text{e}^-/(\text{J} \cdot \text{nm})$
$F$	相对孔径 (光圈数)	无量纲
$g$	系统综合增益	$\text{DN}/\text{e}^-$
$T$	曝光时间	s
$\bar{S}$	感光单元灵敏度因子	无量纲
$A$	感光单元的有效感光面积	$\text{m}^2$
$U$	照度调制因子	无量纲
$D_{\text{offset}}$	偏置响应值	DN
$f(\cdot)$	电路中的非线性调制	-

为了使数据处理更加直观、方便, 本节及后文中将直接对 **RGB 三通道** 原始图像 (以下简称为 raw 图像<sup>†</sup>) 进行操作。为避免在色彩插值的过程中引入颜色失真, 本论文中采用传感器彩色滤色片阵列下采样 (color filter array binning) 的方式将单通道 CFA 图像转换为

\* 数字响应值实际上是无量纲的数值, 但为了表述方便, 本论文中人为地为其加上“DN”这一量纲。

<sup>†</sup> 在一些文献中, raw 图像特指直接由传感器输出的未经任何处理的 CFA 图像。本论文中采用一种更为笼统的说法, 即将系统误差修正前后、正式送入 ISP 之前的图像统称为 raw 图像。

三通道 raw 图像\*。以具有“RGGB”样式 Bayer 阵列的图像传感器为例，本论文中将单通道 CFA 图像中的  $2 \times 2$  像素单元（quad）转换为三通道 RGB raw 图像中的单个像素，其中 RGB 图像中的 R、B 通道响应值与  $2 \times 2$  单元中位于 R、B 滤色片下方感光单元的原始响应值相等，而 G 通道响应值则为  $2 \times 2$  单元中两个位于 G 滤色片下方感光单元的原始响应值的算数平均值。该转换过程如图 3-3 所示。这种四像素合一的彩色滤色片阵列下采样方法以牺牲图像空间分辨率为代价，最大程度保留了传感器所记录的原始响应值信息。在一些文献中，这种下采样后的三通道 raw 图像也被称为非 CFA 图像（non-CFA image）<sup>[115]</sup>。

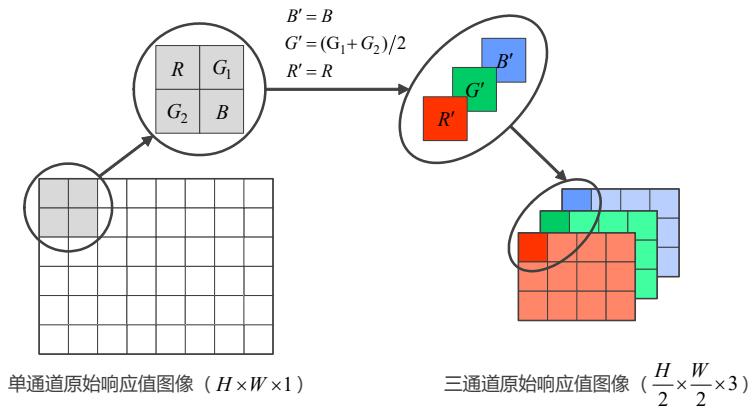


图 3-3 传感器彩色滤色片阵列下采样（CFA binning）

本论文使用 Nikon D3x 单反相机与 Sony ILCE7 微单相机作为测试相机，其中部分与本论文研究内容相关的硬件参数如表 3-2 所示。

在将单通道 CFA 图像转换为三通道 raw 图像后，式 (3.15) 中的各参数对于不同通道的感光单元而言各自独立，因此有

$$\begin{cases} D_r = \text{round} \left[ f \left( \frac{\pi}{4F^2} g_r T \bar{S}AU_r \int_{\lambda} L(\lambda) t_r(\lambda) q(\lambda) d\lambda + D_{\text{offset},r} \right) \right] \\ D_g = \text{round} \left[ f \left( \frac{\pi}{4F^2} g_g T \bar{S}AU_g \int_{\lambda} L(\lambda) t_g(\lambda) q(\lambda) d\lambda + D_{\text{offset},g} \right) \right] \\ D_b = \text{round} \left[ f \left( \frac{\pi}{4F^2} g_b T \bar{S}AU_b \int_{\lambda} L(\lambda) t_b(\lambda) q(\lambda) d\lambda + D_{\text{offset},b} \right) \right] \end{cases} \quad (3.16)$$

\* Sigma 公司所使用的 Foveon 感光技术允许单个感光单元同时侦测长、中、短三种波段的入射光能量，因此无需使用色彩插值模块进行空间插值。本小节中所采用的彩色滤色片阵列下采样方法仅适用于使用单层感光技术的图像传感器。

表 3-2 本章实验中所使用的两台测试相机相关参数。传感器分辨率一栏括号中的数值表示经彩色滤色片阵列下采样（binning）后三通道 raw 图像的实际空间分辨率

相机型号	Nikon D3x	Sony ILCE7
传感器类型	CMOS	CMOS
传感器尺寸	36mm×24mm	36mm×24mm
传感器分辨率	6080×4044 (3040×2022×3)	6024×4024 (3012×2012×3)
原始数据位深	14-bit	12-bit
彩色滤色片阵列格式	Bayer RGGB	Bayer RGGB
镜头型号	Nikkor AF-S 24-120 f/4.0 ED VR	Sonnar T* FE 55 f/1.8 ZA
实验中使用光圈数	f/4.0	f/4.0

## 3.2 数码相机的系统误差修正

为了准确地对光信号与数字信号间的转化模型进行描述，除了考虑原理性的物理过程之外，也有必要对其中可能存在的误差进行分析与修正。理论上，在光源与被摄物体表面保持稳定的情况下，根据式 (3.15)，不同帧中的同一像素，或同一帧中的相邻像素应该具有完全相同的原始响应值\*。然而在实际情况中，以上光信号→电信号→数字信号的过程必然伴随着随机噪声与系统误差的引入。

在对图像传感器的噪声进行建模之前需要将式 (3.15) 中的响应值构成模型扩展到整个传感器平面上，因此有

$$D(i, j) = \text{round} \left[ \frac{\pi T}{4F^2} g(i, j) \bar{S}(i, j) A(i, j) U(i, j) \int_{\lambda} L(\lambda) t(i, j, \lambda) q(i, j, \lambda) d\lambda + D_{\text{offset}}(i, j) \right] \quad (3.17)$$

式中  $(i, j)$  表示某个感光单元在传感器平面上的位置坐标。为了计算方便，已令  $f(x) = x$ 。由于本小节中所涉及的噪声模型以及系统误差修正方法对于三通道中的任一通道均成立，因此为简洁起见，暂且省略表示颜色通道的下标 “ $k$ ”。

\* 相比于物体表面纹理的空间周期，照度调制因子  $U(i, j)$  的变化十分缓慢，因此可认为相邻的像素具有相同的  $U$  值。

### 3.2.1 图像传感器的噪声模型

具体来说，数码相机将光信号转换为数字信号的过程中主要包含像素响应非均匀性、热噪声、散粒噪声、固定模型噪声、读出噪声以及量化误差等六类随机噪声或系统误差。

#### 3.2.1.1 像素响应非均匀性

由于制造加工过程中不确定因素的存在，传感器各个感光单元之间在有效感光面积、光电转换效率、非均匀性、滤色片光谱透过率等属性上必然存在一定的差异。这种像素之间差异通常被称为像素响应非均匀性（pixel response non-uniformity, PRNU）。假设各感光单元在  $q(i, j, \lambda)$  与  $t(i, j, \lambda)$  上仅存在整体强度上的误差而不存在波长相关的误差，那么对于位于  $(i, j)$  处的感光单元，可以使用非均匀性分布函数

$$K(i, j) = \frac{g(i, j)}{g_0} \frac{\bar{S}(i, j)}{\bar{S}_0} \frac{A(i, j)}{A_0} \frac{t(i, j)}{t_0} \frac{q(i, j)}{q_0} \quad (3.18)$$

来表征各分项误差组合后的整体误差<sup>\*</sup>，式中  $g_0$  表示一个无任何制造误差的理想感光单元所对应的  $g$  值（综合增益）， $\bar{S}_0$ 、 $A_0$ 、 $t_0$ 、 $q_0$  同理。

将式 (3.18) 代入式 (3.17)，有

$$D(i, j) = \text{round} [g_0 K(i, j) U(i, j) I_0 + D_{\text{offset}}(i, j)] \quad (3.19)$$

式中

$$I_0 = \frac{\pi T}{4F^2} \bar{S}_0 A_0 \int_{\lambda} L(\lambda) t_0(\lambda) q_0(\lambda) d\lambda \quad (3.20)$$

表示一个理想的感光单元理论上释放出的电子数。

基于当前的半导体制造工艺， $K$  的数值通常仅在单位强度附近很小的一段区间内波动 [109]，因此有

$$\bar{E}(K) = 1, \quad \text{var}(K) \ll 1, \quad (3.21)$$

式中  $\bar{E}(\cdot)$  与  $\text{var}(\cdot)$  分别表示在空间维度上计算均值 (mean) 与样本方差 (sample variance)，下同。

---

<sup>\*</sup>  $K(i, j)$  与  $U(i, j)$  具有十分类似的性质——它们均体现了相机系统的空间非均匀性，且都可以通过标定的方法预先确定。然而， $U(i, j)$  所描述的空间非均匀性属于成像系统本质上的一种“客观缺陷”，而  $K(i, j)$  描述的则是传感器在制造过程中由于加工误差而导致的“主观缺陷”。此外，在第 4.2.1 小节中还将看到， $U(i, j)$  除了受到光学系统设计的影响之外，还会受到入射能量光谱分布的影响，因此必须根据拍摄内容单独进行标定。

显然，像素响应非均匀性分布函数  $K(i, j)$  属于设备相关的系统误差，其不随外界因素变化而变化，因此可通过标定实验预先获得<sup>[116]</sup>。

### 3.2.1.2 热噪声

光电转换器件的热效应将产生一定数量的服从泊松分布的自由电子<sup>[113]</sup>，这些自由电子与光电效应所激发的光生电子组合后共同从感光单元中释放出来，从而使信号出现一定程度的波动，因而此类噪声也被称为热噪声（thermal noise） $N_{Th}$ 。热噪声的强度仅仅取决于曝光时间  $T$  以及环境温度，而与光电效应产生的电子数  $I$  无关<sup>[117]</sup>。

热噪声是一随机变量，根据泊松分布的性质，有

$$E(N_{Th}) = \text{var}(N_{Th}) = \mu_{Th}, \quad (3.22)$$

式中  $\mu_{Th}$  表示一个与曝光时间及环境温度有关的变量， $E(\cdot)$  与  $\text{var}(\cdot)$  分别表示计算数学期望（expectation）与方差（variance），下同。

考虑热噪声后，式 (3.19) 需改写为

$$D(i, j) = \text{round} \left\{ g_0 [K(i, j)U(i, j)I_0 + N_{Th}(i, j)] + D_{\text{offset}}(i, j) \right\} \quad (3.23)$$

### 3.2.1.3 散粒噪声

根据量子理论，从感光单元中释放出的电子在数量上存在一个随机的涨落，该涨落即为散粒噪声（photon shot noise） $N_S$ <sup>[118]</sup>。散粒噪声是由大量单个事件的统计不确定性引起的，其中主要包括输入光子散粒噪声、光生电流散粒噪声与热效应散粒噪声<sup>[119]</sup>。散粒噪声的存在使得感光单元最终释放出的有效电子数服从泊松分布： $(I^* + N_S) \sim \text{Poisson}(I^*)$ ，其中  $I^* = KUI_0 + N_{Th}$  表示在不考虑散粒噪声的情况下该感光单元应释放出的电子个数。

根据泊松分布的性质，有  $E(I^* + N_S) = \text{var}(I^* + N_S) = I^*$ ，因此，对于散粒噪声  $N_S$  自身而言，有

$$E(N_S) = 0, \quad \text{var}(N_S) = I^* = KUI_0 + N_{Th} \quad (3.24)$$

考虑散粒噪声后，式 (3.23) 需改写为

$$D(i, j) = \text{round} \left\{ g_0 [K(i, j)U(i, j)I_0 + N_{Th}(i, j) + N_S(i, j)] + D_{\text{offset}}(i, j) \right\} \quad (3.25)$$

### 3.2.1.4 固定模式噪声

对于 CMOS 图像传感器来说，其各个感光单元所对应的电路中有可能存在一定的噪声，此类噪声通常被称为固定模式噪声（fixed pattern noise, FPN） $N_{FP}$ 。固定模式噪声与

像素响应非均匀性类似，均属于系统误差，两者的区别在于，固定模式噪声是由电路的制造误差引起的，与光电效应无关，且通常呈现出一定的空间分布规律<sup>[113]</sup>。

考虑固定模式噪声后，式 (3.25) 需改写为

$$D(i, j) = \text{round} \left\{ g_0 [K(i, j)U(i, j)I_0 + N_{Th}(i, j) + N_S(i, j) + N_{FP}(i, j)] + D_{\text{offset}}(i, j) \right\} \quad (3.26)$$

### 3.2.1.5 读出噪声

信号放大单元把模拟电压信号进行放大的过程中会引入一定的读出噪声（readout noise） $N_R$ 。读出噪声的数学期望为零<sup>[109]</sup>，其方差与放大单元的增益系数 $g$  线性正相关。

考虑读出噪声后，式 (3.26) 需改写为

$$D(i, j) = \text{round} \left\{ g_0 [K(i, j)U(i, j)I_0 + N_{Th}(i, j) + N_S(i, j) + N_{FP}(i, j) + N_R(i, j)] + D_{\text{offset}}(i, j) \right\} \quad (3.27)$$

### 3.2.1.6 量化误差

由于模数转换单元输出的数字信号需要以整数的形式存储于寄存器中，因此这一连续信号转化为离散信号的过程中必然会产生一定的量化误差（quantization error） $N_Q$ 。若进入模数转换单元的模拟电压信号等概率地分布于实数轴上，则根据式 (3.14)，量化误差将服从区间 $[-\frac{1}{2}, \frac{1}{2}]$  内的均匀分布： $N_Q \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2})$ ，因此有

$$E(N_Q) = 0, \quad \text{var}(N_Q) = \frac{1}{12} \quad (3.28)$$

考虑量化误差后，式 (3.27) 需改写为

$$D(i, j) = g_0 [K(i, j)U(i, j)I_0 + N_{Th}(i, j) + N_S(i, j) + N_{FP}(i, j) + N_R(i, j)] + D_{\text{offset}}(i, j) + N_Q(i, j) \quad (3.29)$$

式 (3.29) 即为考虑系统噪声与误差后的数码相机原始响应值构成模型。

显然，由于热噪声、散粒噪声、读出噪声、量化误差均为随机变量，因此式 (3.29) 中的原始响应值 $D$  也为一随机变量。

表 3-3 总结了上述六种类型的系统误差或噪声的统计特性。

表 3-3 图像传感器的系统误差与噪声的统计特性

噪声来源	符号	分布类型	期望/均值	方差/样本方差
像素响应非均匀性	$K$	—	$\bar{\mu}_K = 1$	$\sigma_K^2 \ll 1$
固定模式噪声	$N_{FP}$	—	$\bar{\mu}_{FP}$ 未知	$\sigma_{FP}^2$ 未知
热噪声	$N_{Th}$	泊松分布	$\mu_{Th}$ 未知	$\sigma_{Th}^2$ 未知
散粒噪声	$N_S$	泊松分布	$\mu_S = 0$	$\sigma_S^2 = KUI_0 + N_{Th}$
读出噪声	$N_R$	未知	$\mu_R = 0$	$\sigma_R^2$ 未知
量化误差	$N_Q$	均匀分布	$\mu_Q = 0$	$\sigma_Q^2 = \frac{1}{12}$

### 3.2.2 数码相机系统误差的标定与修正

在与颜色科学相关的图像处理应用中，通常会采用空间滤波的方法以获得不受随机噪声影响的原始响应值。例如，在数码相机的色度特征化实验中，通过在  $N \times N$  的窗口内对色块区域取均值，从而获得相机对该色块的平均数字响应值<sup>[27]</sup>。基于此，本小节将重点关注三类均值非零的系统误差（像素响应非均匀性分布函数  $K$ 、固定模式噪声  $N_{FP}$  以及响应值偏置量  $D_{\text{offset}}$ ）的标定与修正过程，而对于热噪声、散粒噪声、读出噪声以及量化误差这三类随机噪声，仅给出其参数估计方案而并不尝试进行修正\*。

#### 3.2.2.1 噪声的分解

根据各类噪声的统计特性，式 (3.29) 中的原始响应值构成模型可分解为两部分：

$$D(i, j) = \mu(i, j) + N(i, j) \quad (3.30)$$

式中  $\mu(i, j)$  表示其数学期望， $N(i, j)$  代表随机噪声，即

$$\begin{cases} \mu(i, j) = g_0 [K(i, j)U(i, j)I_0 + N_{FP}(i, j) + \mu_{Th}] + D_{\text{offset}}(i, j) \\ N(i, j) = g_0 [N_{Th}(i, j) - \mu_{Th} + N_S(i, j) + N_R(i, j)] + N_Q(i, j) \end{cases} \quad (3.31)$$

\* 在计算机视觉与图像处理领域，图像降噪（denoising）通常指的是利用图像中所提供的上下文信息对散粒噪声、读出噪声这两类随机噪声进行抑制，从而复原出无噪声（noise-free）的纯净图像<sup>[120, 121]</sup>。然而这种处理在本质上总是依赖于某些先验知识而非理论模型，因此并不能保证一定能够可靠地抑制噪声。为了尽可能保留传感器所记录的“真实”信息，本论文中并未尝试对此类随机噪声进行抑制。同时，为了与传统的“图像降噪”概念进行区分，本论文中采用了“系统误差修正”这一说法。

显然,  $N(i, j)$  是一个期望为零的随机变量。

进一步地, 根据噪声与电子数  $I$  之间的相关性,  $N(i, j)$  可被分解为电子数相关噪声  $N_I(i, j)$  以及电子数无关噪声  $N_C(i, j)$ :

$$\begin{cases} N_I(i, j) = g_0 N_S(i, j) \\ N_C(i, j) = g_0 [N_{Th}(i, j) + N_R(i, j) - \mu_{Th}] + N_Q(i, j) \end{cases} \quad (3.32)$$

根据式 (3.24) 可知,

$$\begin{aligned} \sigma_I^2(i, j) &= g_0^2 \sigma_S^2(i, j) \\ &= g_0^2 [K(i, j)U(i, j)I_0 + N_{Th}(i, j)] \end{aligned} \quad (3.33)$$

同时, 根据式 (3.22) 与 (3.28) 可知

$$\sigma_C^2 = g_0^2 (\mu_{Th} + \sigma_R^2) + \frac{1}{12} \quad (3.34)$$

式中热噪声与读出噪声均与感光单元所处的位置无关, 因此  $\mu_{Th}$  与  $\sigma_R^2$  可省略位置坐标。

### 3.2.2.2 整体随机噪声的估计

若对于同一场景使用完全相同的参数拍摄  $n_1$  幅图像, 则根据式 (3.30), 有

$$\begin{cases} D_1(i, j) = \mu_1(i, j) + N_1(i, j) \\ \vdots \\ D_{n_1}(i, j) = \mu_{n_1}(i, j) + N_{n_1}(i, j) \end{cases} \quad (3.35)$$

当拍摄的图像数量  $n_1$  足够大时, 利用式 (3.31) 中  $E[N(i, j)] = 0$  这一性质, 可得到关于  $\mu(i, j)$  的一致性估计结果:

$$\begin{aligned} \hat{\mu}(i, j) &\approx \tilde{E}[\mu(i, j)] \\ &= \frac{1}{n_1} \sum_{p=1}^{n_1} \mu_p(i, j) \\ &= \frac{1}{n_1} \sum_{p=1}^{n_1} D_p(i, j) - \frac{1}{n_1} \sum_{p=1}^{n_1} N_p(i, j) \\ &\approx \frac{1}{n_1} \sum_{p=1}^{n_1} D_p(i, j) - E[N(i, j)] \\ &= \frac{1}{n_1} \sum_{p=1}^{n_1} D_p(i, j) \end{aligned} \quad (3.36)$$

式中  $\tilde{E}(\cdot)$  表示在时域中计算均值 (temporal mean), 下同。

此外, 利用  $\tilde{\text{var}}[N(i, j)]$  作为  $\text{var}[N(i, j)]$  的近似<sup>[109]</sup>, 还可得到关于  $\sigma_N^2(i, j)$  的一致性估计:

$$\begin{aligned}
 \hat{\sigma}_N^2(i, j) &\approx \tilde{\text{var}}[N(i, j)] \\
 &= \tilde{\text{var}}[D(i, j)] \\
 &= \frac{1}{n_i-1} \sum_{p=1}^{n_i} [D_p(i, j) - \hat{\mu}(i, j)]^2
 \end{aligned} \tag{3.37}$$

式中  $\tilde{\text{var}}(\cdot)$  表示在时域中计算样本方差 (temporal sample variance), 下同。在本论文实验中,  $n_i = 16$ 。

图 3-4 展示了针对本论文中所使用的两台数码相机, 在不同 ISO 下使用式 (3.37) 计算得到的传感器 G 通道整体随机噪声分布情况, 其中拍摄对象为一均匀白板, 表面亮度根据 ISO 的不同分别设置为  $15\sim240 \text{ cd/m}^2$  不等。

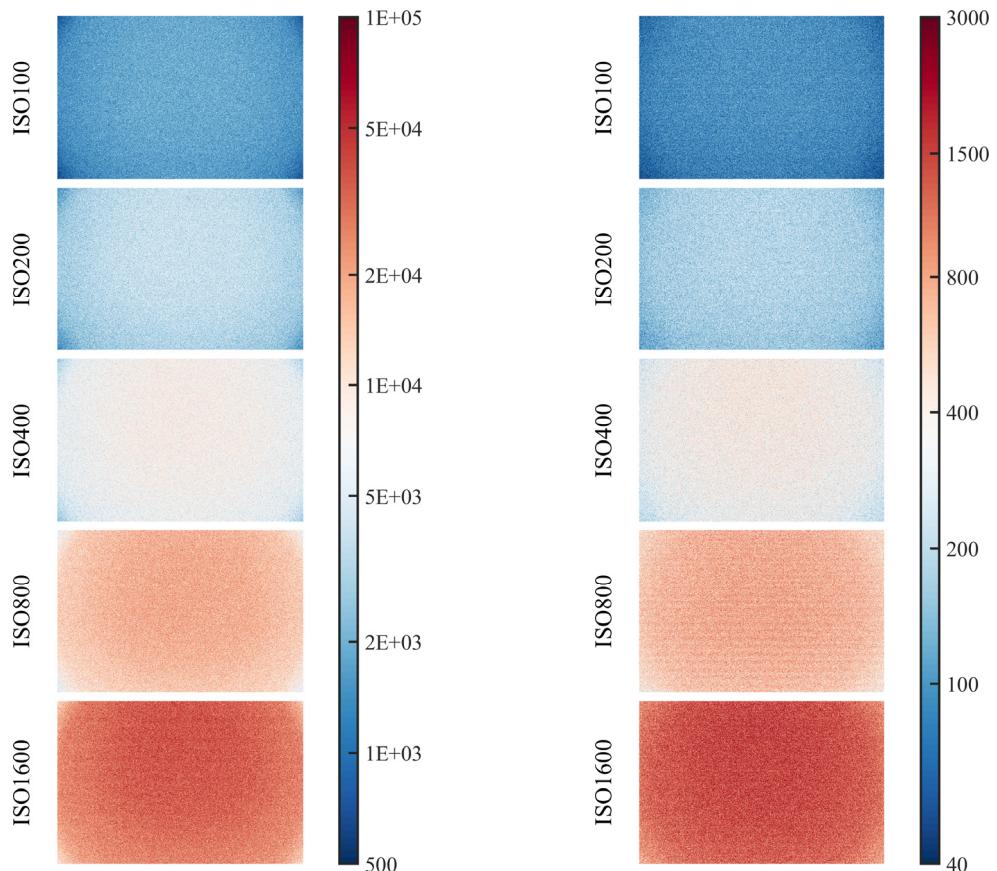


图 3-4 Nikon D3x (左) 与 Sony ILCE7 (右) 在不同 ISO 下的 G 通道整体随机噪声分布 (即  $\hat{\sigma}_N^2(i, j)$ ) 示意图 (曝光时间  $T$  均为  $1/8\text{s}$ ), 其中, Nikon D3x 对应  $b=14$  (14-bit 位深), Sony ILCE7 对应  $b=12$  (12-bit 位深)

### 3.2.2.3 综合增益系数 $g_\theta$ 的估计

由于  $N_I$  与  $N_C$  之间不存在相关性, 因此根据式 (3.33), 整体随机噪声的方差  $\sigma_N^2$  可表示为

$$\begin{aligned}\sigma_N^2(i, j) &= \sigma_I^2(i, j) + \sigma_C^2(i, j) \\ &= g_0^2 [K(i, j)U(i, j)I_0 + N_{Th}(i, j)] + \sigma_C^2\end{aligned}\quad (3.38)$$

针对式 (3.31) 中的  $\mu(i, j)$  以及式 (3.38) 中的  $\sigma_N^2(i, j)$ , 倘若对等号两侧的变量同时计算局部空间均值 (local spatial mean), 则有

$$\begin{aligned}\bar{E}^{w_1} [\mu(i, j)] &= g_0 \left\{ \bar{E}^{w_1} [K(i, j)U(i, j)] I_0 + \bar{E}^{w_1} [N_{FP}(i, j)] + \mu_{Th} \right\} \\ &\quad + \bar{E}^{w_1} [D_{\text{offset}}(i, j)]\end{aligned}\quad (3.39)$$

以及

$$\bar{E}^{w_1} [\sigma_N^2(i, j)] = g_0^2 \left\{ \bar{E}^{w_1} [K(i, j)U(i, j)] I_0 + \bar{E}^{w_1} [N_{Th}(i, j)] \right\} + \sigma_C^2 \quad (3.40)$$

式中  $\bar{E}^{w_1}[X(i, j)]$  以及后文中的  $\bar{E}^{w_1}(i, j)$  均表示以  $(i, j)$  为中心, 在  $w_1 \times w_1$  的窗口内对  $X$  进行空间均值滤波。

一方面, 由于照度调制因子  $U(i, j)$  的空间变化十分缓慢 (详见第 4.2.1 小节), 当  $w_1$  足够小时, 可认为在该窗口内  $U(i, j)$  为一常数  $\bar{U}^{w_1}$ ; 另一方面, 当  $w_1$  足够大时, 即窗口内包含的像素数足够多时,  $K(i, j)$ 、 $N_{Th}(i, j)$  以及  $D_{\text{offset}}(i, j)$  的空间均值也可视为常数, 且根据前文, 有

$$\begin{cases} \bar{E}^{w_1} [K(i, j)] \approx 1 \\ \bar{E}^{w_1} [N_{Th}(i, j)] \approx \mu_{Th}, \quad \text{if } w_1 \text{ is large enough} \\ \bar{E}^{w_1} [D_{\text{offset}}(i, j)] \approx \mu_{\text{offset}} \end{cases} \quad (3.41)$$

式中  $\mu_{\text{offset}}$  表示响应值偏置量的数学期望, 通常可从传感器的参数手册中获知。

将式 (3.41) 代入式 (3.39)、(3.40) 中, 有

$$\begin{cases} \bar{\mu}^{w_1}(i, j) = g_0 \left[ \bar{U}^{w_1}(i, j) I_0 + \bar{N}_{FP}^{w_1}(i, j) + \mu_{Th} \right] + \mu_{\text{offset}} \\ \bar{\sigma}_N^{w_1}(i, j) = g_0^2 \left[ \bar{U}^{w_1}(i, j) I_0 + \mu_{Th} \right] + \sigma_C^2 \end{cases} \quad (3.42)$$

显然,  $\bar{\mu}^{w_1}(i, j)$  与  $\bar{\sigma}_N^{w_1}(i, j)$  之间满足

$$\bar{\sigma}_N^{w_1}(i, j) = g_0 \bar{\mu}^{w_1}(i, j) + \left[ \sigma_C^2 - g_0 \mu_{\text{offset}} - g_0^2 \bar{N}_{FP}^{w_1}(i, j) \right] \quad (3.43)$$

因此, 只要获得足够多的  $\left[ \bar{\mu}^{w_1}(i, j), \bar{\sigma}_N^{w_1}(i, j) \right]$  样本点, 就可以通过线性回归的方式估计出

一条斜率为  $g_0(i, j)$ 、截距为  $\left[ \sigma_C^2 - g_0(i, j)\mu_{\text{offset}} - g_0^2 \bar{N}_{FP}^{w_1}(i, j) \right]$  的直线。实际操作时，可通过改变光源亮度的方式<sup>[122]</sup>获得  $n_2$  组不同的  $D(i, j)$ ，并对式 (3.36) 与 (3.37) 进行空间均值滤波，从而得到  $n_2$  组  $\left[ \bar{\mu}^{w_1}(i, j), \bar{\sigma}_N^{2w_1}(i, j) \right]$  坐标。

当  $n_1$  较大时，使用式 (3.37) 估计出的  $\hat{\sigma}_N^2$  近似服从正态分布<sup>[123]</sup>，因此可以使用最大似然估计对  $g_0(i, j)$  进行估计<sup>[124]</sup>：

$$\hat{g}_0(i, j) = \frac{\bar{XY} - \bar{X}\bar{Y}}{\bar{X}^2 - \bar{X}^2} \quad (3.44)$$

其中

$$\begin{cases} \bar{X} = \frac{1}{n_2} \sum_{q=1}^{n_2} \bar{\mu}_q^{w_1}(i, j) \\ \bar{Y} = \frac{1}{n_2} \sum_{q=1}^{n_2} \bar{\sigma}_N^{2w_1}(i, j) \\ \bar{X}^2 = \frac{1}{n_2} \sum_{q=1}^{n_2} [\bar{\mu}_q^{w_1}(i, j)]^2 \\ \bar{XY} = \frac{1}{n_2} \sum_{q=1}^{n_2} \bar{\mu}_q^{w_1}(i, j) \bar{\sigma}_N^{2w_1}(i, j) \end{cases} \quad (3.45)$$

式中  $\bar{\mu}_q^{w_1}(i, j)$  表示在第  $q$  幅图像（即第  $q$  组光源亮度等级下拍摄的图像，且  $1 \leq q \leq n_2$ ）上对  $\hat{\mu}(i, j)$  取  $w_1 \times w_1$  的窗口并计算局部空间均值， $\bar{\sigma}_N^{2w_1}(i, j)$  同理。

理论上  $g_0$  为一位置无关的全局常量，因此通过对各像素对应的  $\hat{g}_0(i, j)$  计算空间均值的方式获得最终的全局  $g_0$  估计。在本论文实验中， $w_1 = 15$ ， $n_2 = 8$ 。

图 3-5 展示了针对本论文中所使用的两台数码相机，在不同 ISO 下使用式 (3.44) 并计算全局均值后得到的传感器综合增益估计。

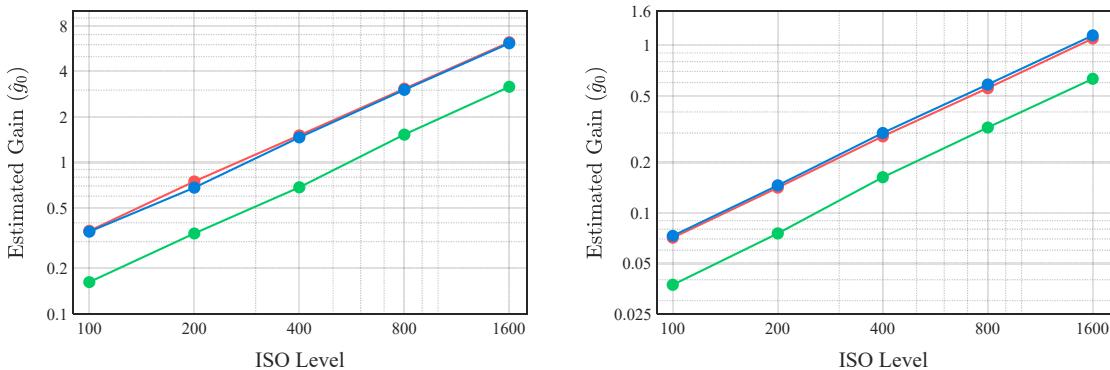


图 3-5 Nikon D3x (左) 与 Sony ILCE7 (右) 在不同 ISO 下的 RGB 三通道综合增益系数

### 3.2.2.4 暗电流噪声的估计

在完全黑暗的条件下使用相机进行拍摄，此时有  $I_0 = 0$ ，故式 (3.30) 变为

$$D_{\text{dark}}(i, j) = \mu_{\text{dark}}(i, j) + N_{\text{dark}}(i, j) \quad (3.46)$$

其中

$$\begin{cases} \mu_{\text{dark}}(i, j) = g_0 N_{FP}(i, j) + g_0 \mu_{Th} + D_{\text{offset}}(i, j) \\ N_{\text{dark}}(i, j) = g_0 [N_{Th}(i, j) + N_{S, \text{dark}}(i, j) + N_R(i, j) - \mu_{Th}] + N_Q(i, j) \end{cases} \quad (3.47)$$

显然， $\mu_{\text{dark}}(i, j)$  对应了传感器在未接收到任何光信号的情况下对应的数字响应值期望，因此通常也被称为数码相机的暗电流 (dark current) \*。

仍然利用  $E[N_{\text{dark}}(i, j)] = 0$  这一性质，在黑暗环境下拍摄  $n_3$  幅图像并在时域上取均值，即可得到关于暗电流噪声  $\mu_{\text{dark}}(i, j)$  的一致性估计：

$$\begin{aligned} \hat{\mu}_{\text{dark}}(i, j) &\approx \tilde{E}[\mu_{\text{dark}}(i, j)] \\ &= \frac{1}{n_3} \sum_{p=1}^{n_3} D_{\text{dark}, p}(i, j) - \frac{1}{n_3} \sum_{p=1}^{n_3} N_{\text{dark}, p}(i, j) \\ &\approx \frac{1}{n_3} \sum_{p=1}^{n_3} D_{\text{dark}, p}(i, j) - E[N_{\text{dark}}(i, j)] \\ &= \frac{1}{n_3} \sum_{p=1}^{n_3} D_{\text{dark}, p}(i, j) \end{aligned} \quad (3.48)$$

本论文实验中， $n_3 = 16$ 。

图 3-6 展示了针对本论文中所使用的两台数码相机，使用式 (3.48) 计算得到的 G 通道暗电流分布情况，其中对于 Nikon D3x 有  $\mu_{\text{offset}} = 0$ ，对于 Sony ILCE7 有  $\mu_{\text{offset}} = 128$ 。

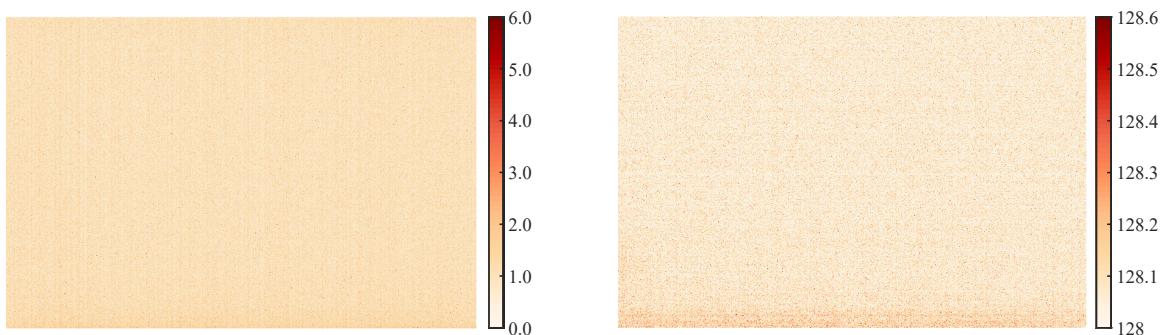


图 3-6 Nikon D3x (左) 与 Sony ILCE7 (右) 在  $T=1/8s$ 、ISO=100 下的 G 通道暗电流噪声  $\mu_{\text{dark}}(i, j)$  分布示意图

\* 一些文献中也称之为黑电平 (black level)。

### 3.2.2.5 像素响应非均匀性的估计

为表达方便起见，我们首先令  $e(i, j) = K(i, j)U(i, j)I_0$ 。若在  $e(i, j)$  内以  $(i, j)$  为中心，在  $w_2 \times w_2$  的窗口内对  $e(i, j)$  计算空间均值，那么一方面，当  $w_2$  足够小时，可认为在该窗口内  $U(i, j)$  为一常数  $\bar{U}^{w_2}$ ，另一方面，当  $w_2$  足够大时，有  $\bar{E}^{w_2}[K(i, j)] \approx 1$ 。因此，只需选取合适的  $w_2$ ，即可得到

$$\bar{e}^{w_2}(i, j) = \bar{E}^{w_2}[e(i, j)] \approx \bar{U}^{w_2}(i, j)I_0 \quad (3.49)$$

显然， $e(i, j)$  与  $\bar{e}^{w_2}(i, j)$  之间满足

$$\frac{e(i, j)}{\bar{e}^{w_2}(i, j)} = \frac{K(i, j)U(i, j)}{\bar{U}^{w_2}(i, j)} \approx K(i, j) \quad (3.50)$$

因此，只需要获得足够多的  $[\bar{e}^{w_2}(i, j), e(i, j)]$  样本点，就可以通过线性回归的方式估计出一条斜率为  $K(i, j)$  且经过原点的直线。实际操作时，可通过改变光源亮度的方式获得  $n_4$  组  $[\bar{e}^{w_2}(i, j), e(i, j)]$  坐标。

利用最小二乘法，有

$$\hat{K}(i, j) = \frac{1}{n_4} \sum_{q=1}^{n_4} \frac{e_q(i, j)}{\bar{e}_q^{w_2}(i, j)} \quad (3.51)$$

式中  $e_q(i, j)$  表示第  $q$  幅图像（即第  $q$  组光源亮度等级下拍摄的图像，且  $1 \leq q \leq n_4$ ）对应的  $e(i, j)$ ， $\bar{e}_q^{w_2}(i, j)$  同理。

为避免拍摄的物体表面存在缺陷或污点，可在每次切换光源时稍微移动被摄物体，这样即使在  $n_4$  组样本点中存在个别异常点，也很容易发现并剔除，从而保证线性回归模型的鲁棒性<sup>[109]</sup>。

下面讨论  $e(i, j)$  的计算。

联立式 (3.30)、(3.31) 可得

$$D(i, j) = g_0[e(i, j) + N_{FP}(i, j) + \mu_{Th}] + D_{\text{offset}}(i, j) + N(i, j) \quad (3.52)$$

再一次地，利用  $E[N(i, j)] = 0$  这一性质，可通过在同一条件下拍摄  $n_5$  幅图像并计算时域均值的方式消去  $N(i, j)$ ：

$$\begin{aligned} \tilde{E}[D(i, j)] &= \frac{1}{n_5} \sum_{p=1}^{n_5} D_p(i, j) \\ &\approx g_0[e(i, j) + N_{FP}(i, j) + \mu_{Th}] + D_{\text{offset}}(i, j) \end{aligned} \quad (3.53)$$

因此，联立式 (3.53) 与 (3.47)，即可得到关于  $e(i, j)$  的估计：

$$\hat{e}(i, j) \approx \frac{\frac{1}{n_5} \sum_{p=1}^{n_5} D_p(i, j) - \hat{\mu}_{\text{dark}}(i, j)}{\hat{g}_0} \quad (3.54)$$

本论文实验中， $n_5 = 16$ ， $n_4 = 8$ ， $w_2 = 15$ 。

图 3-7 展示了针对本论文中所使用的两台数码相机，利用式 (3.51) 计算得到的 G 通道对应的像素响应非均匀性  $\hat{K}(i, j)$  的分布情况。

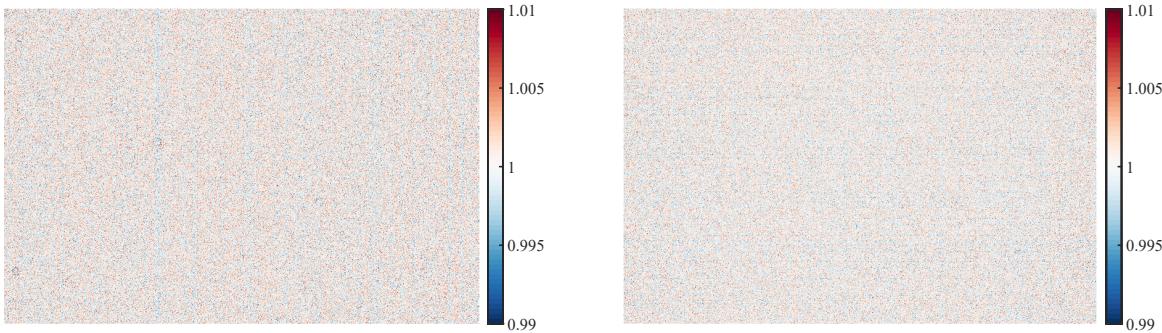


图 3-7 Nikon D3x (左) 与 Sony ILCE7 (右) 的 G 通道像素响应非均匀性  $\hat{K}(i, j)$  分布示意图

### 3.2.2.6 Raw 图像的系统误差修正

将式 (3.48)、(3.51) 中估计出的暗电流噪声  $\hat{\mu}_{\text{dark}}(i, j)$  与像素响应非均匀性分布函数  $\hat{K}(i, j)$  代入式 (3.29) 中，即可得到系统误差修正后的原始响应值（下标“nc”表示“noise correction”）：

$$\begin{aligned} D_{\text{nc}}(i, j) &= \frac{D(i, j) - \hat{\mu}_{\text{dark}}(i, j)}{\hat{K}(i, j)} \\ &= g_0 U(i, j) I_0 + \frac{\epsilon(i, j)}{K(i, j)} \end{aligned} \quad (3.55)$$

式中

$$\epsilon(i, j) = g_0 [N_{Th}(i, j) - \mu_{Th} + N_S(i, j) + N_R(i, j)] + N_Q(i, j) \quad (3.56)$$

表示一个期望为零的随机变量。

图 3-8 中展示了对于 Nikon D3x 拍摄的测试 raw 图像使用式 (3.55) 进行系统误差修正后结果，其中相对误差  $D_{\text{diff}} = 100\% \times \left| \frac{D_{\text{nc}} - D}{D_{\text{nc}}} \right|$ 。



图 3-8 对 Nikon D3x 拍摄的测试 raw 图像（左）进行系统误差修正后的图像（中）以及两者 G 通道间的相对误差（右）

由于随机噪声  $\epsilon(i, j)$  的期望为零，所以可以通过拍摄多幅 raw 图像并取时域平均的方式去除该随机噪声。对于单张 raw 图像中较为平缓的区域（例如颜色单一、均匀的色块表面等），也可采用计算空间均值的方式去除随机噪声：当计算空间均值的窗口  $w_3 \times w_3$  足够大且各像素相互独立时，该窗口内各像素的均值与该点的数学期望近似，因此有

$$\begin{aligned} \bar{D}_{\text{nc}}^{w_3}(i, j) &= g_0 \bar{U}^{w_3}(i, j) I_0 + \bar{E}^{w_3} \left[ \frac{\epsilon(i, j)}{K(i, j)} \right] \\ &\approx g_0 \bar{U}^{w_3}(i, j) I_0 + \frac{E[\epsilon(i, j)]}{\bar{E}^{w_3}[K(i, j)]} \\ &= g_0 \bar{U}^{w_3}(i, j) I_0 \\ &= D_0(i, j) \end{aligned} \quad (3.57)$$

式中  $I_0$  即为传感器  $(i, j)$  处感光单元在无噪声的理想情况下对应的原始响应值。

联立式 (3.55)、(3.57) 可得

$$D_0(i, j) \approx \bar{D}_{\text{nc}}^{w_3}(i, j) = \bar{E}^{w_3} \left[ \frac{D(i, j) - \hat{\mu}_{\text{dark}}(i, j)}{\hat{K}(i, j)} \right] \quad (3.58)$$

该式即为针对单张 raw 图像的噪声修正模型。

### 3.3 基于光谱灵敏度函数估计的数码相机原始响应值预测模型

假设数码相机 raw 图像中的三类系统误差均已被修正，随机噪声已通过计算局部均值的方式被抑制，且空间非均匀性  $U(i, j)$  也已得到补偿\*，则联立式 (3.16)、(3.20) 以及 (3.29)

\* 具体实施方法请参见第 4.2 节。照度调制函数  $U(i, j)$  为 raw 图像引入了不必要的空间非均匀性，在一定程度上也可被视为系统误差，因此在 ISP 流程中必须给予修正。

有

$$\begin{cases} D_r(i,j) = \frac{\pi}{4F^2} g_{r,0} T \bar{S}_0 A_0 \int_{\lambda} L(\lambda) t_{r,0}(\lambda) q_0(\lambda) d\lambda \\ D_g(i,j) = \frac{\pi}{4F^2} g_{g,0} T \bar{S}_0 A_0 \int_{\lambda} L(\lambda) t_{g,0}(\lambda) q_0(\lambda) d\lambda \\ D_b(i,j) = \frac{\pi}{4F^2} g_{b,0} T \bar{S}_0 A_0 \int_{\lambda} L(\lambda) t_{b,0}(\lambda) q_0(\lambda) d\lambda \end{cases} \quad (3.59)$$

显然，对于一台数码相机，若能够获得其  $\bar{S}_0$ 、 $A_0$ 、 $t_{k,0}(\lambda)$ 、 $q_0(\lambda)$  等特性参数的准确估计，则对于任意具有光谱辐亮度  $L(\lambda)$  的待测物体表面，均可使用式 (3.59) 对其原始响应值  $D_k$  进行预测，从而构建一台“虚拟相机”并由此对真实世界中的成像过程进行仿真。

在不更换镜头的情况下，对于一个相机系统而言，其光谱透过率函数与光电转换效率函数的乘积可视为一整体，即  $s_{k,0}(\lambda) = t_{k,0}(\lambda) q_0(\lambda)$ ，其量纲与  $q_0(\lambda)$  一致。简化起见，本节中仅针对传感器上某一感光单元或局部若干个感光单元（局部区域内  $U(i,j)$  可视为常数）进行分析\*，因此式 (3.59) 中的位置坐标  $(i,j)$  可被省略，此时有

$$\begin{cases} D_r = g_r T \kappa \int_{\lambda} L(\lambda) s_r(\lambda) d\lambda \\ D_g = g_g T \kappa \int_{\lambda} L(\lambda) s_g(\lambda) d\lambda \\ D_b = g_b T \kappa \int_{\lambda} L(\lambda) s_b(\lambda) d\lambda \end{cases} \quad (3.60)$$

式中  $\kappa = \frac{\pi}{4F^2} \bar{S}_0 A_0$  为设备相关的常量†， $s_k(\lambda)$  表示 R、G、B 三种颜色通道感光单元所对应的光谱透过率与光电转换效率函数的乘积。本节中，在不引起歧义的情况下，所有下标“0”均将被省略。

由于  $s_k(\lambda)$ ,  $k \in \{r,g,b\}$  反映了整个相机系统的光谱特性，因此我们将其称为**数码相机的光谱灵敏度函数** (camera spectral sensitivity functions)。本节中将通过对数码相机的光谱灵敏度函数进行估计从而建立被摄物体的光谱辐亮度  $L(\lambda)$  与其原始响应值  $D_k$  之间的映射关系，进而实现数码相机原始响应值预测模型的构建。通过将该预测模型输出的响应值与真实捕获的图像中经误差修正后的原始响应值进行比较，即可对其预测精度进行定量评估。

\* 若需要对传感器上所有感光单元进行全局分析，只需将传感器平面进行若干个子区域划分，并将本节中的局部分析方法应用于各子区域之中，同时使用插值算法确保子区域间具有平滑过渡即可。

† 改变相对孔径  $F$  时会相应地改变相机系统的某些光学性质，因此本论文实验中采用固定的相对孔径大小 ( $F=4$ )。

### 3.3.1 原始响应值构成模型中的非线性效应

在第 0 节关于传感器噪声标定的实验中，出于计算方便的考虑，我们暂时忽略了相机系统中的非线性效应，即令  $f(x) = x$ 。然而，为了得到更加准确的原始响应值预测，有必要在响应值构成模型中将这一非线性调制纳入考虑。

为了确定非线性调制函数  $f(x)$  的具体表达形式，我们首先进行了两组预实验。

#### 3.3.1.1 预实验 A：黑场下原始响应值与拍摄参数的关系

理想情况下，通常认为曝光时间与 ISO 感光度（即模拟增益）对于原始响应值幅值大小的作用是等效的，换言之，若不考虑噪声，那么在延长一倍曝光时间的同时固定 ISO 不变，与在提高一倍 ISO 的同时固定曝光时间不变，对于原始响应值的影响应该是完全等价的。然而，在实际的拍摄过程中，尤其是光信号极为微弱的场景中，由于光电转换过程中非线性的存在，这一等效性并非严格成立。

为验证这一猜想，在第一组预实验中我们首先测试了暗场下原始响应值与曝光时间  $T$  之间的关系，如图 3-9 所示。其中，暗场响应值表示在暗场图像（暗室中盖上镜头盖时捕获的图像）中央  $50 \times 50\text{px}$  区域内计算得到的算数平均值。显然，从图 3-9 中可以得知，在 ISO 感光度固定的情况下，调节曝光时间对于暗场响应值几乎没有影响，由此可以推断，原始响应值构成模型中非线性函数的截距项应与曝光时间无关。

同时，我们也测试了暗场响应值与 ISO 感光度之间的关系，如图 3-10 所示。可以看出，即使经过了系统误差修正，两台测试相机的暗场响应值与 ISO 感光度之间仍然存在一定的相关性。由此可以推断，原始响应值构成模型中非线性函数的截距项应与 ISO 感光度相关。

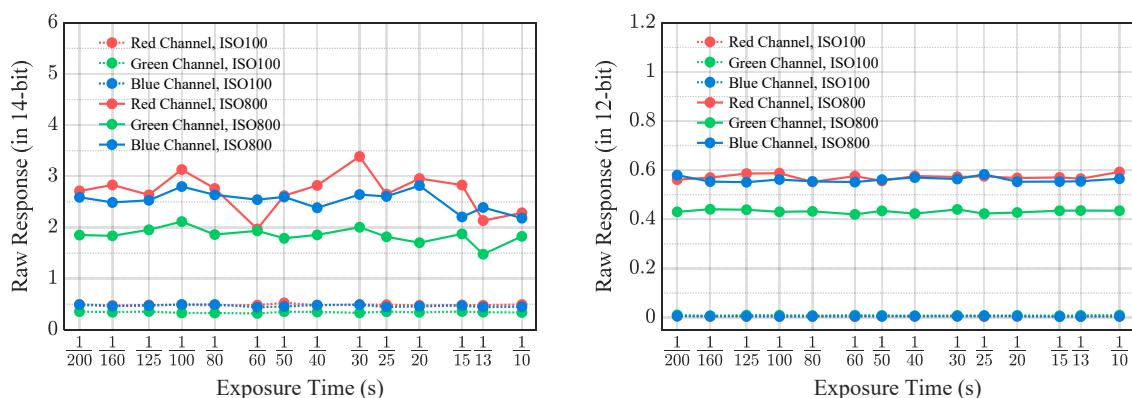


图 3-9 Nikon D3x（左）与 Sony ILCE7（右）在不同曝光时间下的暗场响应值

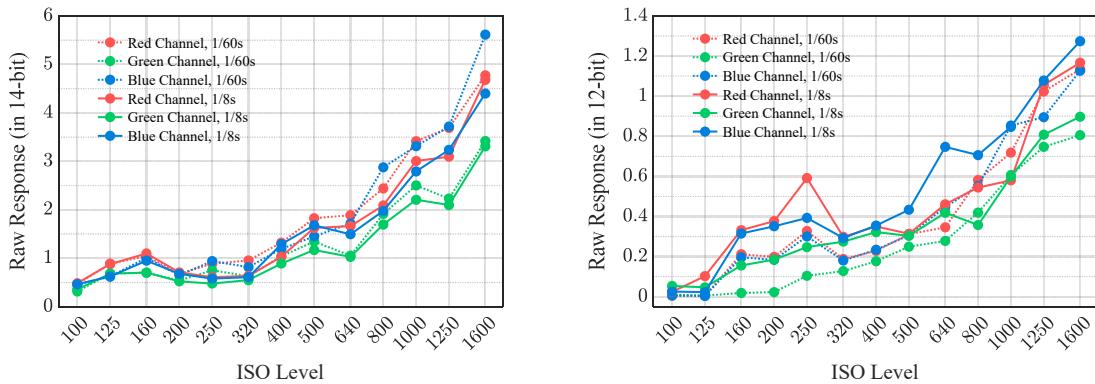


图 3-10 Nikon D3x (左) 与 Sony ILCE7 (右) 在不同 ISO 感光度下的暗场响应值

### 3.3.1.2 预实验 B: 曝光时间与 ISO 感光度的等效性

在第二组预实验中，我们测试了正常光照条件下曝光时间与 ISO 感光度之间的等效性，即，固定  $T$  与 ISO 的乘积，考察在不同拍摄参数组合下的原始响应值的变化趋势。

该组预实验的结果如图 3-11 所示，其中横坐标中带有“\*”的拍摄参数组合表示该组参数对应的原始响应值已预先经过人为的幅值调整以确保曝光不变性（即曝光时间与 ISO 感光度的乘积的不变性）。例如，使用 ISO160、 $T=1/13s$  参数进行拍摄时，其曝光时间与 ISO 的乘积（ $160/13=12.308$ ）是 ISO100、 $T=1/8s$  参数下乘积（ $100/8=12.5$ ）的 0.985 倍，因此对于 ISO160、 $T=1/13s$  参数下获得的响应值，需要将其乘上系数（ $1/0.985=1.016$ ）以满足曝光不变性。此时，经过调整后的曝光参数等价于 ISO160、 $T=1/12.8s$ 。

从图 3-11 中可以看出，当原始响应值足够大时，曝光时间与 ISO 感光度之间近似具有等效性，此时原始响应值构成模型中的常数项（非线性函数的截距项）可忽略不计。

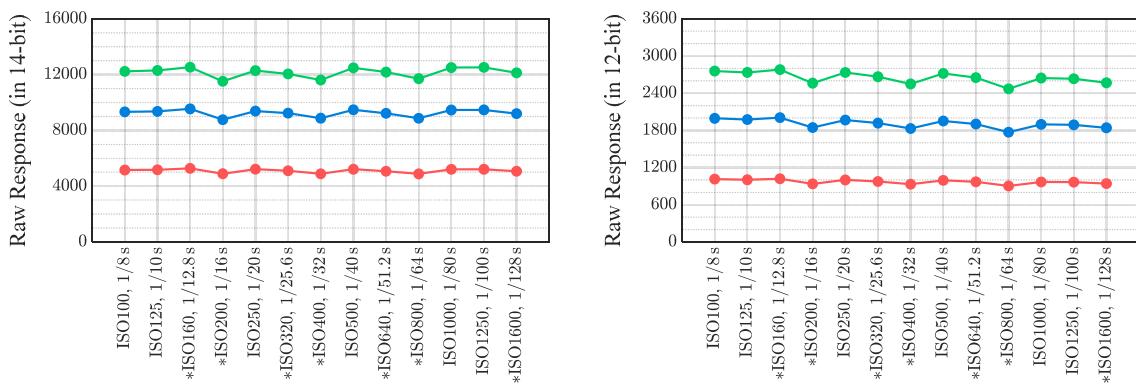


图 3-11 Nikon D3x (左) 与 Sony ILCE7 (右) 在不同拍摄参数组合下的响应值

基于以上两组预实验中的结论，并考虑到模拟电路中的非线性因素，我们对式 (3.60)

进行如下修正：

$$\begin{cases} D_r = \left[ g_r \left( T \kappa \int_{\lambda} L(\lambda) s_r(\lambda) d\lambda + \alpha_r \right) \right]^{\gamma_r} + \beta_r \\ D_g = \left[ g_g \left( T \kappa \int_{\lambda} L(\lambda) s_g(\lambda) d\lambda + \alpha_g \right) \right]^{\gamma_g} + \beta_g \\ D_b = \left[ g_b \left( T \kappa \int_{\lambda} L(\lambda) s_b(\lambda) d\lambda + \alpha_b \right) \right]^{\gamma_b} + \beta_b \end{cases} \quad (3.61)$$

式中参数  $\alpha_k$ 、 $\beta_k$ 、 $\gamma_k$ ,  $k \in \{r, g, b\}$  共同决定了传感器非线性调制的形式，且根据[110]，应有  $\gamma_k \approx 1$ 。

显然， $\left[ (g_k \alpha_k)^{\gamma_k} + \beta_k \right]$ ,  $k \in \{r, g, b\}$  即对应了原始响应值构成模型中的截距项。在第 3.3.3 小节中，我们将以该截距的数值作为优化任务中的约束之一对光谱灵敏度函数进行估计。

### 3.3.2 通道串扰效应

为获取相机的光谱灵敏度函数，传统的物理测量方法通常利用单色仪或窄带滤光片生成目标波段内各个波长上的准单色光，并使用相机对这些准单色光分别进行成像。通过比较相机对于不同准单色光的原始响应值大小，即可得到其光谱灵敏度函数的近似估计。图 3-12 中展示了利用物理测量法进行光谱灵敏度函数估计所需的实验设置。

物理测量法虽然直观且易操作，但是其测量过程中可能引入各种仪器误差，因此若使用该方法得到的光谱灵敏度函数对原始响应值进行预测，其精度并不能令人满意。此外，物理测量法使用准单色光作为光源，从根本上抑制了传感器相邻感光单元之间可能存在的通道串扰效应，从而导致其预测的响应值与实际拍摄过程中获得的原始响应值之间存在一定的误差。

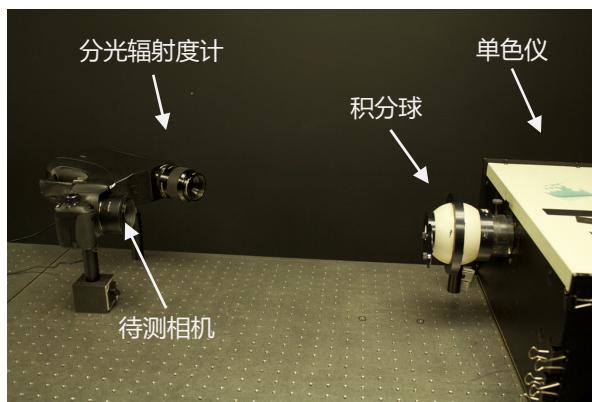


图 3-12 使用物理测量法获取相机光谱灵敏度函数的实验设置 © J. Jiang<sup>[23]</sup>

在式 (3.60) 所描述的响应值构成模型中，我们理想地认为每个感光单元所输出的原始响应值仅仅取决于其对应的光谱灵敏度函数。然而，当入射光线倾斜地照射在传感器某一区域上时，该区域内相邻的感光单元之间并不能严格地保证不存在能量泄漏<sup>[125]</sup>。例如，穿过某一 G 滤色片的部分能量有可能落在与其相邻的 R 感光单元的光敏面上，同理，穿过 R 滤色片的能量也可能并非全部到达 R 感光单元光敏面上，而是部分泄漏到了相邻的其它感光单元内部。因此，此时 R 感光单元对应的原始响应值不再严格地等于  $g_r T \kappa \int_{\lambda} L(\lambda) s_r(\lambda) d\lambda$ ，而是等于

$$g_r T \kappa \int_{\lambda} L(\lambda) [c_r s_r(\lambda) + c_g s_g(\lambda)] d\lambda \quad (3.62)$$

式中  $c_r$  与  $c_g$  分别表示 R、G 感光单元对应的能量占比系数。这种对应着不同颜色通道的相邻感光单元之间的能量泄漏，通常被称为传感器的通道串扰 (inter-channel crosstalk) 效应，如图 3-13 所示。显然，入射光线的倾斜角越大，或相邻感光单元之间的隔离程度越差，则通道串扰效应越严重<sup>[126]</sup>。

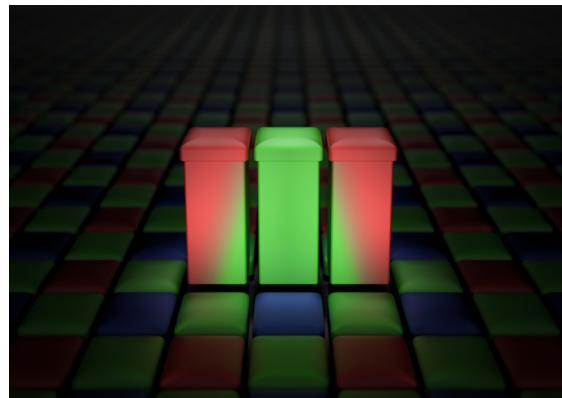


图 3-13 传感器 R、G 感光单元之间的通道串扰效应示意图

使用准单色光对相机系统进行照明时，其通道串扰效应的严重程度与实际拍摄情况下存在较大的差别。不妨考虑这么一种较为特殊的情况：假设某一相机的 R、G 滤色片分别仅允许 550~700nm 以及 450~600nm 波段内的能量透过。为了测量 R 通道光谱灵敏度函数  $s_r(\lambda)$  在 650nm 波长处的幅值，需要使用中心波长位于 650nm 的准单色光作为光源。显然，使用相机对该单色光进行拍摄时，其 R 感光单元上接收的能量将全部来源于透过 R 滤色片的光线，而落在与之相邻的 G 感光单元上的能量将完全被 G 滤色片截去（G 滤色片不允许通过 650nm 处的入射光），因此，即使该传感器存在着通道串扰效应，这部分能量也无法到达 R 感光单元的光敏面上。然而，在实际的拍摄场景中，几乎不存在这种单峰且带

宽极窄的光源。假设将上文中的 650nm 准单色光替换为一个具有双脉冲的窄带光源，其中一个脉冲仍位于 650nm 处，而另一个则位于 500nm 处。此时，若传感器存在通道串扰，则波长为 500nm 的光线将有可能从 G 滤色片中透过并落在 R 感光单元的光敏面上，此时 R 感光单元输出的原始响应值将同时受到 650nm 与 500nm 波长能量的影响，这显然与准单色光照明时的情况有所出入。

对于移动终端（如手机、平板电脑等）中的相机这类出瞳距较短的成像系统，其传感器边缘的感光单元对应了很大的入射光线倾斜角（亦称为主光线夹角，chief ray angle），因此此类设备的通道串扰效应相较于专业相机而言也要更加严重。此时若使用物理测量法获得的光谱灵敏度函数对响应值进行预测，则必然存在一定的误差。

考虑到通道串扰效应对响应值构成模型产生的影响，有必要对式 (3.61) 进一步进行修正：

$$\begin{cases} D_r = \left[ g_r \left( T \kappa \int_{\lambda} L(\lambda) [c_{rr} s_r(\lambda) + c_{gr} s_g(\lambda) + c_{br} s_b(\lambda)] d\lambda + \alpha_r \right) \right]^{\gamma_r} + \beta_r \\ D_g = \left[ g_g \left( T \kappa \int_{\lambda} L(\lambda) [c_{rg} s_r(\lambda) + c_{gg} s_g(\lambda) + c_{bg} s_b(\lambda)] d\lambda + \alpha_g \right) \right]^{\gamma_g} + \beta_g \\ D_b = \left[ g_b \left( T \kappa \int_{\lambda} L(\lambda) [c_{rb} s_r(\lambda) + c_{gb} s_g(\lambda) + c_{bb} s_b(\lambda)] d\lambda + \alpha_b \right) \right]^{\gamma_b} + \beta_b \end{cases} \quad (3.63)$$

式中系数  $c_{kk'}$ ,  $k, k' \in \{r, g, b\}$  表示从  $k$  通道感光单元泄漏至  $k'$  通道感光单元的能量占比。

### 3.3.3 原始响应值构成模型中的参数估计

在式 (3.63) 的原始响应值构成模型中，已知量包括：

- 测试物体的光谱辐亮度  $L(\lambda)$ （可用分光辐射度计测得）；
- 曝光时间  $T$ ；
- 系统综合增益  $g_k$ （获取方法请参见第 3.2.2.3 小节）；
- 经系统误差修正与随机噪声抑制后的原始响应值  $D_k$ 。

而待估计的参数包括：

- 光谱灵敏度函数  $s_k(\lambda)$ ；
- 设备相关常数  $\kappa$ ；
- 非线性调制因子  $\alpha_k$ 、 $\beta_k$ 、 $\gamma_k$ ；
- 通道串扰因子  $c_{kk'}$ 。

由于  $s_k(\lambda)$ 、 $\kappa$  以及  $c_{kk'}$  三者具有乘积等效性，必须对其施加一定的约束以得到具体的数值解。不失一般性地，令  $s_k(\lambda)$  表示数码相机的相对光谱灵敏度函数，因此有\*

$$\max_{\substack{k \in \{r,g,b\} \\ 380\text{nm} \leq \lambda \leq 780\text{nm}}} \{s_k(\lambda)\} = 1 \quad (3.64)$$

同时，令通道串扰因子不大于 1，有

$$\max_{\substack{k \in \{r,g,b\} \\ k' \in \{r,g,b\}}} \{c_{kk'}\} = 1 \quad (3.65)$$

出于计算方便的考虑，我们将式 (3.63) 转换为矩阵运算的形式：

$$\mathbf{D} = \mathbf{g} \left[ T\kappa(\Delta\lambda) \mathbf{L} \cdot \mathbf{S}^T \cdot \mathbf{C} + \mathbf{a} \right]^\gamma + \mathbf{\beta} \quad (3.66)$$

式中：

- $\mathbf{D}$  表示  $1 \times 3$  (1 行 3 列，下同) 的原始响应值行向量，即

$$\mathbf{D} = [D_r, D_g, D_b] \quad (3.67)$$

由于不同型号的相机往往采用不同的存储位深，因此在对原始响应值进行进一步的操作之前有必要先对其进行归一化处理：

$$\mathbf{D} \leftarrow \frac{\mathbf{D}}{2^b - 1} \quad (3.68)$$

式中  $b$  代表位深。

- $\mathbf{L}$  表示  $1 \times N$  的光谱辐亮度行向量，即

$$\mathbf{L} = [L(380), \dots, L(780)] \quad (3.69)$$

式中  $N$  代表分光辐射度计的采样波长数。

- $\mathbf{S}^T$  表示  $N \times 3$  的相机光谱灵敏度函数矩阵，其每一列对应一个颜色通道的相机光谱灵敏度函数，即

$$\mathbf{S}^T = \begin{bmatrix} s_r(380) & s_g(380) & s_b(380) \\ \vdots & \vdots & \vdots \\ s_r(780) & s_g(780) & s_b(780) \end{bmatrix} \quad (3.70)$$

- $\mathbf{C}$  表示  $3 \times 3$  的通道串扰矩阵，其  $p$  行  $p'$  列的元素数值代表第  $p$  个颜色通道对第  $p'$  个颜色通道产生的通道串扰因子，即

\* 注意：该最大值对应的是三通道光谱灵敏度函数的共同最大值，而非每个通道内的局部最大值。

$$\mathbf{C} = \begin{bmatrix} c_{rr} & c_{gr} & c_{br} \\ c_{rg} & c_{gg} & c_{bg} \\ c_{rb} & c_{gb} & c_{bb} \end{bmatrix} \quad (3.71)$$

- $\alpha$ 、 $\beta$ 、 $\gamma$  分别表示三通道非线性调制因子  $\alpha_k$ 、 $\beta_k$ 、 $\gamma_k$  各自对应的  $1 \times 3$  行向量，即

$$\begin{cases} \boldsymbol{\alpha} = [\alpha_r, \alpha_g, \alpha_b] \\ \boldsymbol{\beta} = [\beta_r, \beta_g, \beta_b] \\ \boldsymbol{\gamma} = [\gamma_r, \gamma_g, \gamma_b] \end{cases} \quad (3.72)$$

- $\mathbf{g}$  表示由三通道综合增益系数构成的  $1 \times 3$  行向量，即

$$\mathbf{g} = [g_r, g_g, g_b] \quad (3.73)$$

- $\Delta\lambda$  代表分光辐射度计的采样波长间隔，由于光谱灵敏度函数已被限制在 380~780nm 内合计共 400nm 的波长范围内，因此有  $(N-1)\Delta\lambda = 400$ 。

需要注意的是，在式 (3.66) 中，除了“.” 符号明确表示矩阵乘法外，其余运算符号均表示逐元素操作 (element-wise operation)。

当使用数码相机作为测量仪器对辐射量进行记录时，通常希望其原始响应值与被摄物体的辐亮度之间尽可能地满足线性关系。因此，一旦获得了式 (3.66) 中的非线性参数以及通道串扰矩阵后，便可对其进行逆变换以获得理想的线性原始响应值  $\mathbf{D}_0$ ：

$$\begin{aligned} \mathbf{D}_0 &= \mathbf{g} T \kappa(\Delta\lambda) \mathbf{L} \cdot \mathbf{S}^T \\ &= \mathbf{g} \left[ \left( \frac{\mathbf{D} - \boldsymbol{\beta}}{\mathbf{g}} \right)^{1/\gamma} - \boldsymbol{\alpha} \right] \cdot \mathbf{C}^{-1} \end{aligned} \quad (3.74)$$

同样地，式 (3.74) 中除了“.” 符号表示矩阵乘法外，其余运算符号均表示逐元素操作。

### 3.3.3.1 损失函数

正式地，我们将关于  $\mathbf{S}$ ,  $\mathbf{C}$ ,  $\kappa$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$  等参数的估计问题表示为

$$(\hat{\mathbf{S}}, \hat{\mathbf{C}}, \hat{\kappa}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \arg \min \left[ \frac{1}{M} \sum_i^M \mathcal{L}(\hat{\mathbf{D}}_i, \mathbf{D}_i) \right] \quad (3.75)$$

式中  $M$  为训练样本的个数， $\mathcal{L}(\cdot, \cdot)$  表示损失函数， $\mathbf{D}_i$  表示第  $i$  个训练样本的真实原始响应值向量（可从 raw 图像中提取），而

$$\hat{\mathbf{D}}_i = \mathbf{g} \left[ T \hat{\kappa}(\Delta\lambda) \mathbf{L}_i \cdot \hat{\mathbf{S}}^T \cdot \hat{\mathbf{C}} + \hat{\boldsymbol{\alpha}} \right]^{\hat{\gamma}} + \hat{\boldsymbol{\beta}} \quad (3.76)$$

则表示将各估计参数代入式 (3.66) 后得到的预测原始响应值向量。

出于计算方便的考虑，一些研究中直接使用原始响应值三维空间中  $\hat{\mathbf{D}}_i$  与  $\mathbf{D}_i$  之间的欧氏距离作为光谱灵敏度函数估计的损失函数<sup>[22, 24, 127-130]</sup>。然而，一方面，数码相机的原始响应值总是设备相关 (device-dependent) 的，在一个设备相关的空间中计算得到的欧氏距离大小并不能客观反映模型性能的优劣；另一方面，本论文中对原始响应值构成模型进行研究的最终目的，并非只是对光谱灵敏度函数  $s_k(\lambda)$  或原始响应值  $D_k$  进行估计，而是希望借助该模型的预测值对 ISP 流程各模块中的参数进行调节，从而提高数码相机的颜色复现精度。

基于这一目标，我们使用色度特征化后响应值之间的 CIEDE2000 色差作为损失函数对原始响应值构成模型中的参数进行端到端 (end-to-end) 的估计，即

$$\mathcal{L}(\hat{\mathbf{D}}_i, \mathbf{D}_i) \triangleq \Delta E_{00} [\mathcal{P}(\hat{\mathbf{D}}_i), \mathcal{P}(\mathbf{D}_i)] \quad (3.77)$$

式中  $\mathcal{P}(\cdot)$  表示色度特征化操作，其具体实现过程将在后文中阐述。

同理，在对原始响应值预测模型的性能进行评估时，我们也使用色度特征化操作  $\mathcal{P}(\cdot)$  将预测得到的原始响应值转换至设备无关颜色空间中，并使用 CIEDE2000 色差作为其预测精度的评价指标。

### 3.3.3.2 参数估计中的色度特征化

在对原始响应值构成模型进行参数估计时，色度特征化的目的仅仅是为了将设备相关的原始响应值变换至设备无关颜色空间中，因此无需区分训练和验证的过程，只需确保对  $\hat{\mathbf{D}}_i$  与  $\mathbf{D}_i$  所使用的色度特征化模型完全一致即可。

在综合考虑了特征化精度与模型复杂度之后，我们选择使用三次项的齐次根式多项式模型 (3rd degree root-polynomial model)<sup>[63]</sup> 作为参数估计过程中的色度特征化模型。三次项根式多项式模型使用一个  $13 \times 3$  颜色空间变换矩阵  $\mathbf{M}$  将原始响应值映射至 CIE1931 XYZ 颜色空间中。该变换矩阵的计算过程简述如下：

- 1) 将每个训练样本的真实原始响应值  $\mathbf{D}_i$  从  $1 \times 3$  行向量扩展为  $1 \times 13$  的行向量  $\tilde{\mathbf{D}}_i$ ：

$$\begin{aligned} \tilde{\mathbf{D}}_i = & [D_{r,i}, D_{g,i}, D_{b,i}, \sqrt{D_{r,i}D_{g,i}}, \sqrt{D_{r,i}D_{b,i}}, \sqrt{D_{g,i}D_{b,i}}, \\ & \sqrt[3]{D_{r,i}D_{g,i}^2}, \sqrt[3]{D_{r,i}D_{b,i}^2}, \sqrt[3]{D_{g,i}D_{b,i}^2}, \\ & \sqrt[3]{D_{g,i}D_{r,i}^2}, \sqrt[3]{D_{b,i}D_{r,i}^2}, \sqrt[3]{D_{b,i}D_{g,i}^2}, \\ & \sqrt[3]{D_{r,i}D_{g,i}D_{b,i}}] \end{aligned} \quad (3.78)$$

式中  $D_{k,i}, k \in \{r, g, b\}$  表示第  $i$  个样本  $k$  颜色通道的真实原始响应值；

2) 将分光辐射度计测得的每个训练样本的光谱辐亮度数据  $L(\lambda)$  作为  $\varphi(\lambda)$  代入式 (2.12) 中，

计算得到该样本在同一光源下真实的 XYZ 三刺激值  $\mathbf{X}_i = [X_i, Y_i, Z_i]$ ；

3) 使用非线性优化方法<sup>\*</sup>计算最佳的变换矩阵  $\mathbf{M}$  以最小化  $M$  个样本上的 CIEDE2000 色差均值：

$$\mathbf{M} = \arg \min \left[ \frac{1}{M} \sum_i^M \Delta E_{00} (\tilde{\mathbf{D}}_i \cdot \mathbf{M}, \mathbf{X}_i) \right] \quad (3.79)$$

由于变换矩阵  $\mathbf{M}$  具有光源依赖性，因此当使用不同的光源对原始响应值预测模型的预测精度进行测试时，也必须分别计算相应的  $\mathbf{M}$  以确保色度特征化的可靠性。

需要说明的是，利用式 (3.79) 计算得到的变换矩阵  $\mathbf{M}$  虽然并不能保证色度特征化后的响应值 ( $\tilde{\mathbf{D}}_i \cdot \mathbf{M}$ ) 与真实的三刺激  $\mathbf{X}_i$  之间严格相等，但由于该变换矩阵将同时作用于  $\hat{\mathbf{D}}_i$  与  $\mathbf{D}_i$ ，因此该色度特征化操作对于两者而言都是“公平”的。

综上，式 (3.75) 中的参数估计问题可改写为

$$(\hat{\mathbf{S}}, \hat{\mathbf{C}}, \hat{\kappa}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \arg \min \left[ \frac{1}{M} \sum_i^M \Delta E_{00} (\tilde{\mathbf{D}}_i \cdot \mathbf{M}, \tilde{\mathbf{D}}_i \cdot \mathbf{M}) \right] \quad (3.80)$$

### 3.3.3.3 参数初始化

本论文选择使用 interior-point 算法对式 (3.80) 中带约束的非线性优化问题进行求解 [131, 132]。与其它的非线性优化算法类似，interior-point 算法需要用户提供一组初始化的参数作为最小值搜索过程的起始点。

在计算初始化参数时，可暂且忽略非线性调制函数以及通道串扰矩阵，即令  $\alpha = \beta = 0$ ， $\gamma = 1$ ， $\mathbf{C} = \mathbf{I}_3$  ( $\mathbf{I}_N$  表示  $N \times N$  的单位矩阵 (identity matrix)，下同)，因此，式 (3.66) 可暂时改写为

$$\frac{\mathbf{D}}{\mathbf{g}} = T \kappa(\Delta \lambda) \mathbf{L} \cdot \mathbf{S}^\top \quad (3.81)$$

式中除号 “-” 表示逐元素除法 ( $\mathbf{D}$  与  $\mathbf{g}$  均为  $1 \times 3$  行向量)。

类似地，可将单个颜色样本的响应值表达式拓展至  $M$  个样本上，此时有

<sup>\*</sup> 针对 CIEDE2000 色差的非线性优化方法将在第 4.4.1 小节中进行介绍。

$$\frac{\mathbf{D}'}{\mathbf{g}'} = \kappa(\Delta\lambda) \mathbf{T} \cdot \mathbf{L}' \cdot \mathbf{S}^\top \quad (3.82)$$

式中：

- $\mathbf{D}'$  表示  $M \times 3$  的原始响应值矩阵，每一行对应一个样本的三通道原始响应值，即

$$\mathbf{D}' = \begin{bmatrix} D_{1,r} & D_{1,g} & D_{1,b} \\ \vdots & \vdots & \vdots \\ D_{M,r} & D_{M,g} & D_{M,b} \end{bmatrix} \quad (3.83)$$

- $\mathbf{T}$  表示由各个样本曝光时间构成  $M \times M$  对角矩阵，即

$$\mathbf{T} = \begin{bmatrix} T_1 & & \\ & \ddots & \\ & & T_M \end{bmatrix} \quad (3.84)$$

- $\mathbf{L}'$  表示  $M \times N$  的光谱辐亮度矩阵，其中每一行对应一个样本的光谱辐亮度数据，即

$$(3.85)$$

- $\mathbf{g}'$  表示由各个样本的三通道综合增益系数构成的  $M \times 3$  矩阵，即

$$\mathbf{g}' = \begin{bmatrix} g_{1,r} & g_{1,g} & g_{1,b} \\ \vdots & \vdots & \vdots \\ g_{M,r} & g_{M,g} & g_{M,b} \end{bmatrix} \quad (3.86)$$

- 除号 “-” 仍然表示逐元素除法（ $\mathbf{D}'$  与  $\mathbf{g}'$  均为  $M \times 3$  矩阵）。

由于实测的光谱辐亮度  $\mathbf{L}'$  中噪声的存在，直接使用伪逆法（Moore–Penrose pseudo-inverse）或奇异值分解（SVD）计算得到的光谱灵敏度函数  $\mathbf{S}^\top$  将出现严重的振荡<sup>[133]</sup>。为此，需要在求解式 (3.82) 的过程中加入正则化（regularization）项，以确保最终得到的光谱灵敏度函数满足一定的平滑约束。

对式 (3.82) 进行变形并对其各通道进行独立分析，有

$$(\Delta\lambda) \mathbf{T} \cdot \text{diag}^{-1}\left(\frac{\mathbf{D}'_k}{\mathbf{g}'_k}\right) \cdot \mathbf{L}' \cdot (\kappa \mathbf{S}_k^\top) = \mathbf{1} \quad (3.87)$$

式中下标 “ $k$ ” 表示仅考虑  $k$  通道的元素（即仅考虑矩阵的其中一列），此时  $\mathbf{D}'_k$ 、 $\mathbf{g}'_k$  均为  $M \times 1$  列向量， $\mathbf{S}_k^\top$  为  $N \times 1$  列向量； $\text{diag}(\cdot)$  表示将向量转化为对角矩阵；等号右侧的  $\mathbf{1}$  表示一个元素全为 1 的  $M \times 1$  列向量。由于  $\mathbf{S}_k^\top$  与  $\kappa$  具有乘积等效性，我们暂且将其乘积  $\kappa \mathbf{S}_k^\top$  视为一个整体。

同时，对  $\kappa \mathbf{S}_k^T$  添加平滑约束，有

$$\mathcal{R} \cdot (\kappa \mathbf{S}_k^T) = \mathbf{0} \quad (3.88)$$

式中

$$\mathcal{R} = \begin{bmatrix} -1 & 2 & -1 \\ & -1 & 2 & -1 \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 & -1 \end{bmatrix} \quad (3.89)$$

表示  $(N-2) \times N$  正则化矩阵，等号右侧的  $\mathbf{0}$  表示一个元素全为 0 的  $(N-2) \times 1$  列向量。

合并式 (3.87) 与 (3.88)，可得到带有正则化项的  $k$  通道光谱灵敏度函数  $\kappa \mathbf{S}_k^T$  的最小二乘表达式：

$$\begin{bmatrix} (\Delta\lambda) \mathbf{T} \cdot \text{diag}^{-1} \left( \frac{\mathbf{D}'_k}{\mathbf{g}'_k} \right) \cdot \mathbf{L}' \\ \lambda \mathcal{R} \end{bmatrix} \cdot (\kappa \mathbf{S}_k^T) = \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix} \quad (3.90)$$

式中  $\lambda$  表示正则化参数，合适的  $\lambda$  值应使得最终的解在拟合残差 (residual fitting error)

$$\left\| \left[ (\Delta\lambda) \mathbf{T} \cdot \mathbf{L}' \cdot (\kappa \mathbf{S}_k^T) - \frac{\mathbf{D}'_k}{\mathbf{g}'_k} \right] \middle/ \frac{\mathbf{D}'_k}{\mathbf{g}'_k} \right\|_2 \quad (3.91)$$

与正则化误差 (regularization error)

$$\left\| \lambda \mathcal{R} \cdot (\kappa \mathbf{S}_k^T) \right\|_2 \quad (3.92)$$

之间取得良好的平衡。

对于大部分需要引入正则化约束的病态问题 (ill-posed problem)，若将不同  $\lambda$  对应的拟合残差作为自变量、正则化误差作为因变量在 log-log 平面上进行描绘，则其形成的轨迹将近似于一条“L”型曲线 (L-curve) [134]。这条曲线为计算合适的正则化参数  $\lambda$  提供了一种便捷的图形化分析方法：合适的  $\lambda$  值通常位于该“L”型曲线的拐点附近，因此只需要对该曲线的拐点进行定位，便可以得到合适的正则化参数 [134]。图 3-14 展示了针对本论文中所使用的两台数码相机，使用 L-curve 方法得到的最佳正则化参数  $\lambda$ 。

由于现实中的相机光谱灵敏度函数在各个波长上的数值总是大等于零的，因此在处理式 (3.90) 的最小二乘问题时，还需要为  $\kappa \mathbf{S}_k^T$  添加非负约束 [135]。

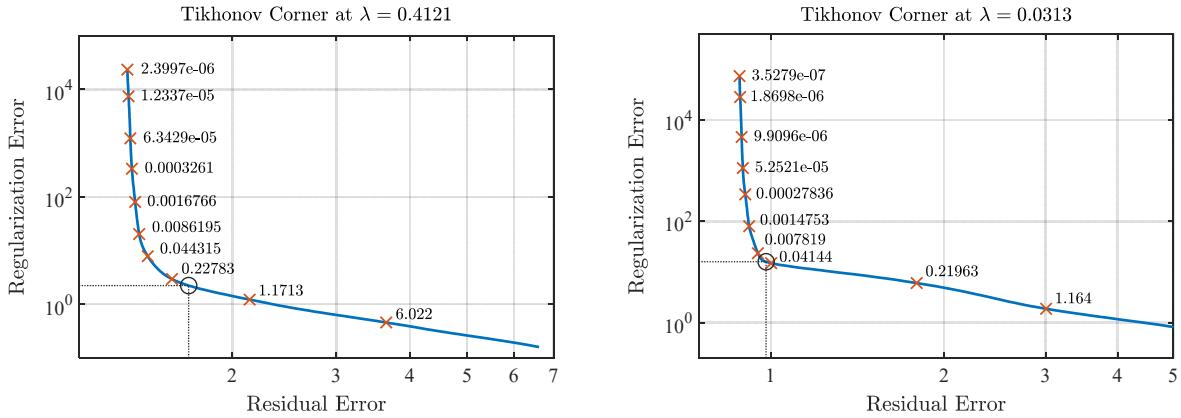


图 3-14 Nikon D3x (左) 与 Sony ILCE7 (右) R 通道对应的 L-curve, 其中位于曲线拐点处的黑色圆圈代表最佳正则化参数

图 3-15 展示了针对本论文中所使用的两台相机, 使用有效集非负最小二乘法 (active set non-negative least squares) [136] 解得的三通道光谱灵敏度函数  $\kappa S_k(\lambda)$ ,  $k \in \{r, g, b\}$ <sup>\*</sup>。该组光谱灵敏度函数也将作为初始值参与到后续的非线性优化过程中。

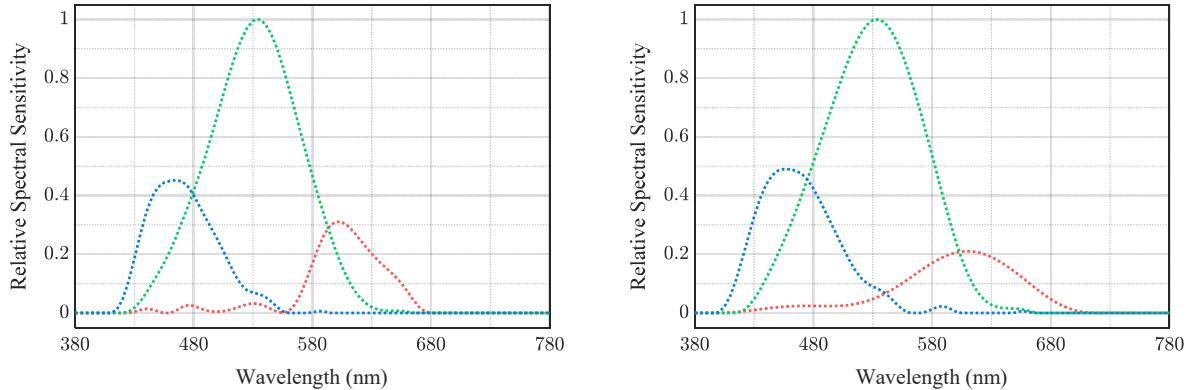


图 3-15 Nikon D3x (左) 与 Sony ILCE7 (右) 对应的 (归一化) 光谱灵敏度函数初始值

在得到光谱灵敏度函数的初始值后, 可使用非线性拟合的方法对非线性参数  $\alpha$ 、 $\beta$ 、 $\gamma$  的初始值进行计算。对于  $k$  通道的非线性参数  $\alpha_k$ 、 $\beta_k$ 、 $\gamma_k$ , 本论文采用 trust region reflective 算法 [137] 对形如

$$y = g(x + \alpha_k)^{\gamma_k} + \beta_k \quad (3.93)$$

\* 此处解得的光谱灵敏度函数  $\kappa S_k^T$  的唯一作用仅仅是为 interior-point 算法提供所需的初始值, 因此不必将  $\kappa$  与  $S_k^T$  进行分离。图 3-15 中为了展示方便, 特意对光谱灵敏度函数做了归一化处理, 即  $s_k(\lambda) = \kappa S_k(\lambda) / \max_{k \in \{r, g, b\}} \{\kappa S_k(\lambda)\}$ 。

的曲线进行最小二乘拟合，其中  $x = (\Delta\lambda)\mathbf{T} \cdot \mathbf{L}' \cdot (\kappa\mathbf{S}_k^T)$  表示使用光谱灵敏度函数初始值（即式 (3.90) 的解）预测得到的线性原始响应值， $y = \mathbf{D}'_k$  则表示真实的原始响应值。与求解光谱灵敏度函数初始值的过程类似，该曲线拟合过程也将对三通道独立进行。

表 3-4 中列出了针对本论文中所使用的两台数码相机，采用最小二乘拟合得到的三通道非线性参数。

表 3-4 使用最小二乘曲线拟合得到的原始响应值构成模型的初始化非线性参数

相机型号	Nikon D3x	Sony ILCE7
$\alpha_r, \alpha_g, \alpha_b$ 数值	-0.0104, -0.0330, -0.0081	-0.0009, -0.2823, -0.0450
$\beta_r, \beta_g, \beta_b$ 数值	0.0141, 0.0155, 0.0164	0.0035, 0.0235, 0.0076
$\gamma_r, \gamma_g, \gamma_b$ 数值	1.0167, 1.0307, 1.0411	1.0510, 1.0139, 1.0255

### 3.3.3.4 非线性参数优化

在得到关于  $\kappa\mathbf{S}_k^T$ 、 $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{\gamma}$ 、 $\mathbf{C}$  的初始化参数后，我们使用 interior-point 算法对式 (3.80) 进行非线性优化<sup>[131, 132]</sup>：

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \begin{cases} \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b} \\ \mathbf{x}_{lb} \leq \mathbf{x} \leq \mathbf{x}_{ub} \end{cases} \quad (3.94)$$

式中， $f(\cdot)$  表示计算平均色差（参见式 (3.80)）， $\mathbf{x}$  表示一个由  $\kappa\mathbf{S}_k^T, k \in \{r, g, b\}$ 、 $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{\gamma}$ 、 $\mathbf{C}$  中各元素连结（concatenation）而成的  $(3N+18) \times 1$  列向量，其中，三通道的光谱灵敏度函数共包含  $3N$  个元素，三通道非线性参数与通道串扰矩阵各包含 9 个元素； $\mathbf{A}$  与  $\mathbf{b}$  为线性不等式约束； $\mathbf{x}_{lb}$  与  $\mathbf{x}_{ub}$  分别代表允许的  $\mathbf{x}$  搜索范围的上下限。显然，该非线性优化过程等价于在一个  $(3N+18)$  维的空间中寻找初始值附近满足一定约束条件的局部最小值。

为避免出现过拟合，我们还为关于  $\mathbf{x}$  的非线性参数优化添加了以下约束：

- 优化后的光谱灵敏度函数应具有一定的平滑度（与参数初始化时加入正则化项的作用类似），因此令

$$\mathbf{A} = \begin{bmatrix} \mathcal{R} & & \mathbf{0} \\ & \mathcal{R} & \mathbf{0} \\ & & \mathcal{R} & \mathbf{0} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} T_s \\ \vdots \\ T_s \end{bmatrix} \quad (3.95)$$

式中  $\mathbf{0}$  代表元素全为 0 的  $(N-2) \times 18$  矩阵，因此  $\mathbf{A}$  为  $(3N-6) \times (3N+18)$  矩阵； $\mathbf{b}$  为

$(3N - 6) \times 1$  列向量； $T_s$  表示平滑度阈值，对于本论文中的两台测试相机 Nikon D3x 与 Sony ILCE7，分别有  $T_s = 4$  以及  $T_s = 8$ 。

- 优化后的光谱灵敏度函数在所有波长上的数值最小不应小于零，最大不应超过初始值的 200%；优化后的非线性参数  $\alpha$ 、 $\beta$ 、 $\gamma$  只允许在初始值的  $\pm 200\%$  范围内浮动；优化后的通道串扰矩阵中对角线上的元素应满足  $0.8 \leq c_{kk'} \leq 1$ , ( $k = k'$ )，非对角线上的元素应满足  $0 \leq c_{kk'} \leq 0.1$ , ( $k \neq k'$ )。综上，搜索范围的上下限可分别表示为

$$\left\{ \begin{array}{l} \mathbf{x}_{lb} = [\overbrace{0, \dots, 0}^{3N}, \alpha_0 - 5|\alpha_0|, \beta_0 - 5|\beta_0|, 0.9\gamma_0, \\ \quad 0.8, 0, 0, 0, 0.8, 0, 0, 0, 0, 0.8]^T \\ \mathbf{x}_{ub} = [2(\kappa\mathbf{S}_r)_0, 2(\kappa\mathbf{S}_g)_0, 2(\kappa\mathbf{S}_b)_0, \alpha_0 + 5|\alpha_0|, \beta_0 + 5|\beta_0|, 1.1\gamma_0, \\ \quad 1, 0.1, 0.1, 0.1, 1, 0.1, 0.1, 0.1, 0.1, 1]^T \end{array} \right. \quad (3.96)$$

式中下标“0”表示上一段落中得到初始化参数。

在式 (3.64) 中，我们将  $s_k(\lambda)$  定义为相对光谱灵敏度函数，因此在得到优化后的  $\kappa\mathbf{S}^T$  后，还需要将  $\kappa$  与  $\mathbf{S}^T$  进行分离：

$$\left\{ \begin{array}{l} \kappa = \max(\kappa\mathbf{S}^T) \\ \mathbf{S}_k^T = \frac{(\kappa\mathbf{S}_k^T)}{\kappa} \end{array} \right. \quad (3.97)$$

式中  $\max(\cdot)$  表示计算矩阵中的最大元素。

图 3-16 与表 3-5 展示了优化后的相机光谱灵敏度函数、非线性参数以及通道串扰矩阵。

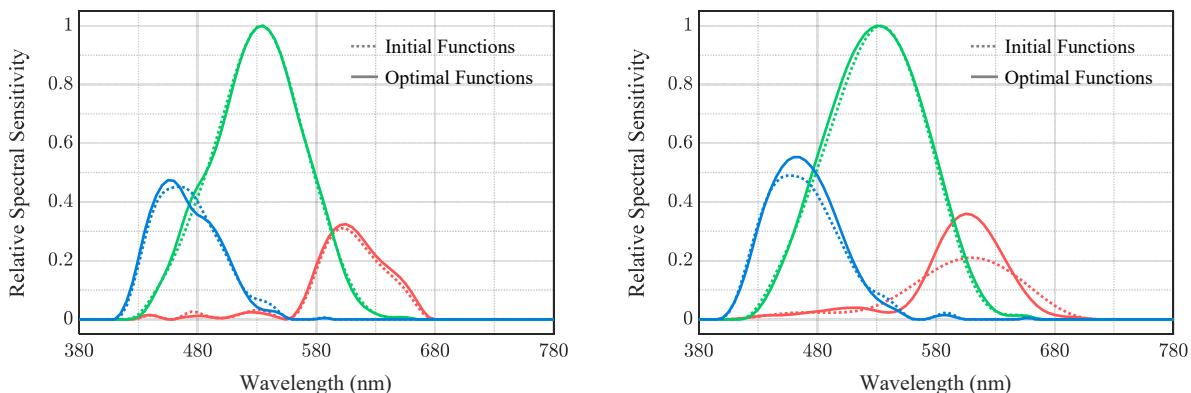


图 3-16 Nikon D3x (左) 与 Sony ILCE7 (右) 经非线性优化后的相对光谱灵敏度函数 (实线) 与初始值 (点线) 的对比结果

在得到最终的预测模型参数后，即可使用式 (3.74) 对图 3-8 中的测试图像进行非线性逆变换。图 3-17 展示了非线性逆变换前后的测试图像颜色外貌及其相对误差。

表 3-5 非线性优化后的原始响应值构成模型参数

相机型号	Nikon D3x	Sony ILCE7
$\kappa$	178.8246	549.7633
$\alpha_r, \alpha_g, \alpha_b$ 数值	-0.0015, -0.0083, -0.0036	0.0007, -0.0310, -0.0028
$\beta_r, \beta_g, \beta_b$ 数值	0.0015, 0.0043, 0.0034	0.0003, 0.0079, 0.0022
$\gamma_r, \gamma_g, \gamma_b$ 数值	1.0581, 1.0471, 1.0414	1.0436, 1.0745, 1.0500
$C$	$\begin{bmatrix} 0.9684 & 0.0534 & 0.0055 \\ 0.0076 & 0.9796 & 0.0130 \\ 0.0178 & 0.0107 & 0.9954 \end{bmatrix}$	$\begin{bmatrix} 0.9231 & 0.0176 & 0.0117 \\ 0.0049 & 0.9735 & 0.0098 \\ 0.0209 & 0.0136 & 0.9909 \end{bmatrix}$



图 3-17 对 Nikon D3x 拍摄的测试 raw 图像（左）进行非线性逆变换后的图像（中）以及两者 G 通道间的相对误差（右）

### 3.3.4 实验设置

#### 3.3.4.1 样本选取

在对原始响应值构成模型中的参数进行估计时，本论文使用 X-Rite Digital SG 色卡中的各个色块作为训练所用的颜色样本。为了使原始响应值预测模型在不同曝光参数下均能对原始响应值进行准确的预测，本论文使用三种不同的曝光参数组合 (ISO100, 1/8s)、(ISO200, 1/15s) 以及 (ISO400, 1/30s) 对颜色样本进行拍摄。同时，为了提高原始响应值预测模型对于不同照明光源的泛化能力，本论文还使用了 X-Rite Macbeth SpectraLight III 灯箱中的 D65 日光模拟器以及 19 通道 LED 平台<sup>[122]</sup>下的 D65 日光模拟器两种光源对样本进行照明。因此，在不同曝光参数以及照明光源组合下，共有  $96 \times 3 \times 2 = 576$  个颜色样本参与

训练过程\*。排除某些原始响应值出现饱和的无效样本，最终的训练样本数  $M(\text{Nikon D3x}) = 496$ ,  $M(\text{Sony ILCE7}) = 473$ 。

为了考察原始响应值预测模型对于不同类型的物体表面以及不同照明光源的泛化能力，在检验阶段，本论文使用不同的色卡以及不同的照明光源对原始响应值预测的准确性进行评估：

- 在第一组检验实验中，使用 X-Rite Macbeth SpectraLight III 灯箱中的 D65 日光模拟器（训练阶段已包含）作为照明光源对 X-Rite Classic 色卡中的各个色块（训练阶段未包含）进行原始响应值预测。该组实验中相机曝光参数以及照明光源的选取均与训练过程相同。X-Rite Classic 色卡包含了 24 个色块，因此该组实验中共有  $24 \times 3 \times 1 = 72$  个检验样本。
- 在第二组检验实验中，使用 X-Rite Macbeth SpectraLight III 灯箱中的 A 光源（训练阶段未包含）作为照明光源对 X-Rite Digital SG 色卡中的各个色块（训练阶段已包含）进行原始响应值预测。该组实验所使用的曝光参数组合与训练阶段略有区别，分别为 (ISO100, 1/10s)、(ISO200, 1/20s) 以及 (ISO400, 1/40s)，因此共有  $96 \times 3 \times 1 = 288$  个检验样本。
- 在第三组检验实验中，同时改变了色卡种类与照明光源，使用 X-Rite Macbeth SpectraLight III 灯箱中的 A 光源（训练阶段未包含）作为照明光源对 X-Rite Classic 色卡中的各个色块（训练阶段未包含）进行响应值预测。该组实验所使用的曝光参数与第二组实验相同，因此共有  $24 \times 3 \times 1 = 72$  个检验样本。

实验中所使用的三种光源(X-Rite Macbeth SpectraLight III 灯箱中的 D65 日光模拟器、A 光源，以及 19 通道 LED 平台下的 D65 日光模拟器) 的光谱功率分布如图 3-18 所示。

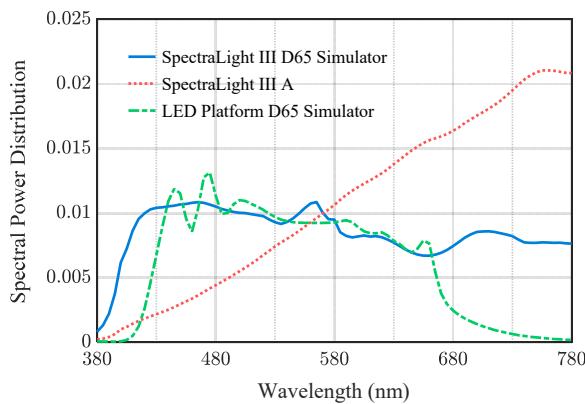


图 3-18 实验中所使用的三种光源的光谱功率分布函数（均归一化至  $Y=1$ ）

\* X-Rite Digital SG 色卡共包含 140 个色块，在参数估计的实验中排除了色卡最外围一圈 44 个重复的中性色色块，因此共有 96 个独立色块参与训练。

### 3.3.4.2 实验设备

在对颜色样本进行拍摄时，光源、样本、相机三者间的距离、相对位置等几何参数都有可能影响到最终记录的原始响应值。为了尽可能排除外界因素的干扰，在拍摄颜色样本的全过程中，我们令光源与相机的绝对位置保持固定，并通过手动移动色卡的方式使得目标色块总是出现在图像的中央区域，如图 3-19 所示。

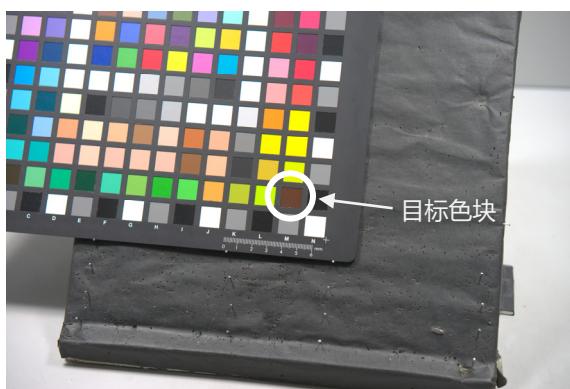


图 3-19 通过手动移动色卡的方式确保目标色块总是出现在图像的中央区域

实验中使用 Konica Minolta CS-2000 分光辐射度计对颜色样本的光谱辐亮度进行测量。CS-2000 能够在 380 至 780nm 的波长范围内以 1nm 的波长间隔进行光谱辐亮度测量，但在实际计算过程中，我们将其测得的光谱数据下采样至 5nm 间隔，因此最终的光谱辐亮度数据共包含 81 个采样点，即  $N=81$ 。测量时，为最大程度消除由于光源不稳定而产生的数据波动，我们令 CS-2000 的每次采样与相机拍摄同步进行。

实验中，CS-2000 与相机均采用 45/0 的测量几何条件（照明光束的光轴方向与样品法线成 45° 角，测量仪器的光轴与与样品法线的夹角小于 10°）<sup>[67]</sup>。由于 CS-2000 与相机的光轴无法同时与目标色块的法线重合，因此我们采用对称放置的方式以使得目标色块在两者光轴方向上的光谱辐亮度尽可能一致，如图 3-20 所示。实际上，相机与分光辐射度计的光轴与目标色块法线的夹角  $\theta$  约 5° 左右。对于 X-Rite Digital SG 色卡，在使用 1° 测量角时，CS-2000 的有效测量范围大约占单个色块总面积的 50% 左右，如图 3-20 中左下角的局部放大图所示（黑色实心圆圈表示光谱辐亮度的有效测量范围）。

在 raw 图像中，我们对每个目标色块中心处  $50 \times 50\text{px}$  的正方形区域（在 55mm 焦距、100cm 拍摄距离的情况下，Digital SG 色卡的单个色块在图像中占据约  $85 \times 85\text{px}$ ）计算均值作为其原始响应值  $\mathbf{D} = [D_r, D_g, D_b]$ （参见式 (3.58)），并依式 (3.68) 进行归一化。

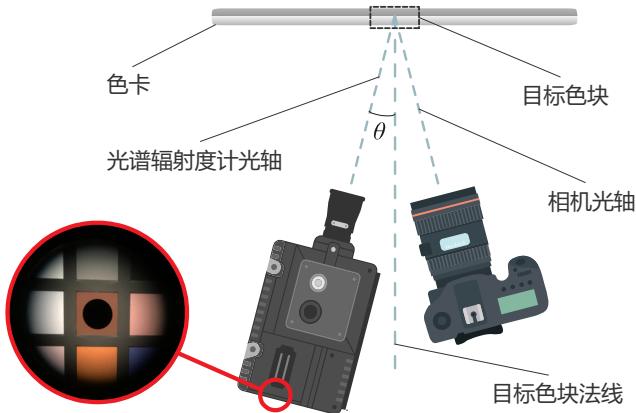


图 3-20 相机拍摄与分光辐射度计测量的几何条件

### 3.3.5 结果与分析

为了对不同条件下的成像过程进行准确的仿真，原始响应值预测模型应具有足够的预测精度与泛化能力。针对每一组检验实验，我们对以下五种原始响应值预测模型的预测结果进行比较：

- PCA：主成分分析（principal component analysis）算法对一组包含 40 台不同品牌型号相机实测数据的光谱灵敏度函数数据库<sup>[23]</sup>进行主成分分析，并将分解后的基向量作为线性基底对待测相机光谱灵敏度函数的各个权重系数进行拟合<sup>[23]</sup>。待估计出相机的光谱灵敏度函数后，利用公式

$$\hat{\mathbf{D}} = \mathbf{g}^T \hat{\kappa}(\Delta\lambda) \mathbf{L} \cdot \hat{\mathbf{S}}^T \quad (3.98)$$

对目标颜色样本的原始响应值进行预测。

- RBFN：径向基函数网络（Radial Basis Functions Network）算法对同样的一组光谱灵敏度函数数据库进行奇异值分解从而获取一组类高斯径向基函数，并将其作为线性基底对待测相机光谱灵敏度函数的各个权重系数进行拟合<sup>[138]</sup>。待估计出相机的光谱灵敏度函数后，同样使用式 (3.98) 对目标颜色样本的原始响应值进行预测。
- PInv：伪逆（Pseudo-Inverse）算法将式 (3.90) 的非负解（即图 3-15 中绘制的曲线）直接作为光谱灵敏度函数的估计值，使用式 (3.98) 对目标颜色样本的原始响应值进行预测。
- Nonlinear：非线性算法将式 (3.93) 的解（即表 3-4 中的数值）直接作为成像系统的非线性参数（未考虑通道串扰效应，即  $\hat{\mathbf{C}} = \mathbf{I}_3$ ），使用式 (3.76) 对目标颜色样本的原始响应值进行预测。
- Optimal：基于优化后的光谱灵敏度函数（即图 3-16 中的实线曲线）与非线性参数（即表 3-5 中的数值），使用式 (3.76) 对目标颜色样本的原始响应值进行预测。

三组检验实验中原始响应值预测精度如图 3-21 ~ 图 3-23 所示, 其中误差线 (error bar) 表示 95% 置信区间。需要说明的是, 在每一组检验实验中, 色度特征化所使用的颜色空间变换矩阵 (参见式 (3.79)) 均是针对当前的目标样本进行优化得到的。该做法的目的在于尽可能地减小不同光源下相机 RGB 颜色空间至 CIE1931XYZ 颜色空间的转换误差。

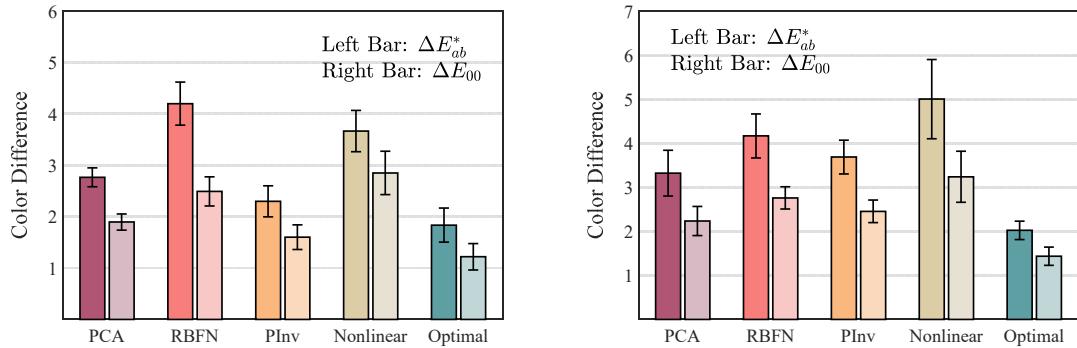


图 3-21 第一组检验实验 (SpectraLight III 灯箱 D65 日光模拟器照明 X-Rite Classic 色卡) 中 Nikon D3x (左) 与 Sony ILCE7 (右) 的原始响应值预测精度

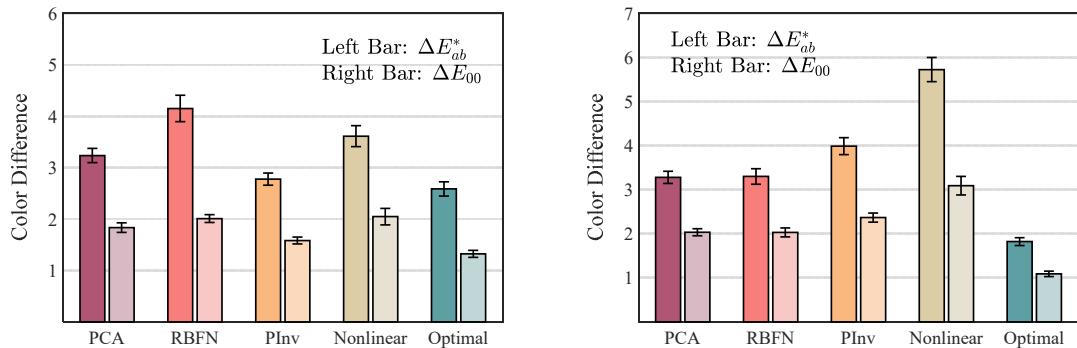


图 3-22 第二组检验实验 (SpectraLight III 灯箱 A 光源照明 X-Rite Digital SG 色卡) 中 Nikon D3x (左) 与 Sony ILCE7 (右) 的原始响应值预测精度

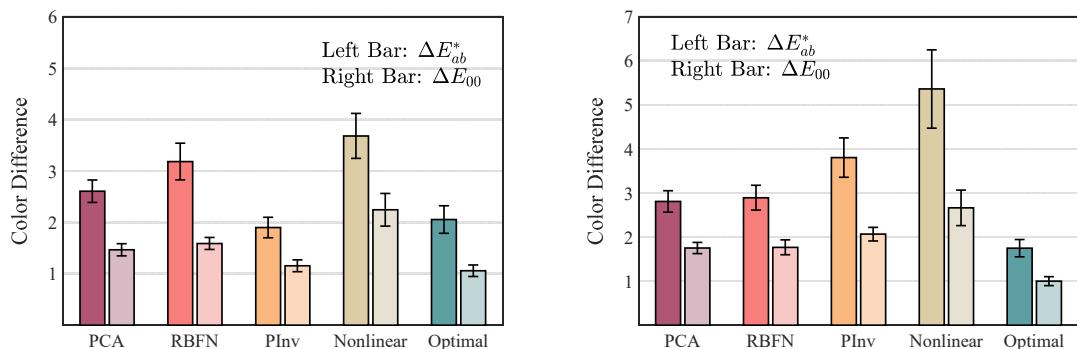


图 3-23 第三组检验实验 (SpectraLight III 灯箱 A 光源照明 X-Rite Classic 色卡) 中 Nikon D3x (左) 与 Sony ILCE7 (右) 的原始响应值预测精度

从图 3-21 ~ 图 3-23 中的结果可以看出，得益于通道串扰矩阵以及非线性参数的引入，本论文中提出的参数估计方法相较于 PCA、RBFN 以及 PInv 三类线性方法在响应值预测精度上具有显著的优势。同时，相较于仅引入非线性参数但未进行任何优化的 Nonlinear 方法，本论文中采用的端到端参数优化方法对于原始响应值估计精度同样起到了明显的改善。此外，图中结果还表明了本论文中提出的原始响应值预测模型对于不同类型的物体表面以及不同种类的照明光源具有较好的泛化性能。在所有的检验实验中，除了 Nikon D3x 对于 SpectraLight III 灯箱 A 光源下 X-Rite Digital SG 色卡的预测精度  $\text{mean}(\Delta E_{ab}^*) = 2.54$  之外，其余颜色样本预测值的平均 CIELAB 色差均小于恰可察觉差(just-noticeable difference, JND)  $\Delta E_{ab}^* \approx 2.3$ <sup>[139]</sup>，由此证明了该原始响应值预测模型能够对不同拍摄条件下的数码相机原始响应值提供可靠的预测。

图 3-24 中展示了使用 PCA、BRFN 以及 Optimal 三种方法估计出的相机光谱灵敏度函数对比结果。需要说明的是，由于 Jiang 等人所提供的相机光谱灵敏度函数数据库<sup>[23]</sup>仅提供了 400~720nm 波长范围内的有效数据，因此图 3-24 中使用 PCA 与 RBFN 算法估计出的光谱灵敏度函数也只包含该波长范围内的数据。

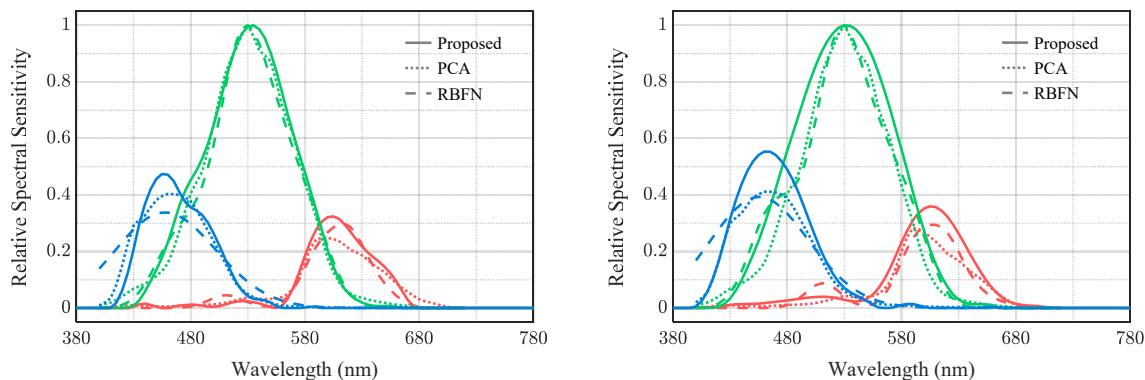


图 3-24 使用 PCA<sup>[23]</sup>、RBFN<sup>[138]</sup>以及本论文所提出的方法估计出的 Nikon D3x（左）与 Sony ILCE7（右）的相对光谱灵敏度函数

### 3.4 本章小结

根据信号载体的不同，本论文将数码相机原始响应值构成的过程分为三个阶段：第一阶段中相机镜头模组通过成像的方式对空间中的辐射能量进行收集，从而将场景中的光信号传递至传感器平面；第二阶段中传感器中的感光元件利用光电转换效应完成光信号至电荷电信号的转换；第三阶段中相机内的读出电路将该电荷信号转换为数字信号，并以原始

数字响应值的形式组成 CFA 图像。本章分别对这三个阶段的信号传输过程进行了数学建模，并最终给出数码相机将场景光谱辐亮度信息转换为原始数字响应值这一物理过程的完整推导。

对于实际的数字成像系统，使用响应值构成模型预测得到的原始响应值与 raw 图像中真实存在的原始响应值之间必然存在一定的误差。这一误差来源主要包括传感器制造过程中的系统误差、模数转换过程中引入的量化误差以及信号传输过程中无法避免的随机误差。为了对原始响应值构成模型进行更加准确的描述，需要对不同种类的误差进行定量表示，并对 raw 图像进行相应的修正。本章基于 CMOS 传感器的硬件特性，对数字成像系统中不同种类的误差来源进行了分析，并对其在信号传输中引入的噪声进行了数学建模。利用本章提出的图像传感器噪声参数估计方法，我们能够在实验室环境下对数字成像设备中的系统误差进行标定与修正，由此最大程度地抑制噪声对于 raw 图像中原始响应值的干扰。

为了构建一台与真实的数码相机具有完全相同响应特性的“虚拟相机”，需要对响应值构成模型中的设备相关参数进行估计，包括光电转换函数、增益系数、系统的非线性等。为此，本章提出了基于预测色差最小化的响应值构成模型参数估计方法。通过对一定数量的训练样本进行拍摄，能够在无需任何标定仪器辅助的情况下完成响应值构成模型中的参数估计任务。在得到模型参数之后，即可通过仿真的手段对任意光源、任意拍摄对象、任意拍摄参数下对应的原始响应值进行预测，从而获得一台与原相机具有十分类似的响应特性的虚拟相机。利用该虚拟相机，我们能够以极高的效率为 ISP 中各模块的标定实验提供大量的训练数据以及优化依据。

## 第四章 数码相机图像信号处理算法研究

### 4.1 图像信号处理流程概述

针对颜色准确复现这一目标，典型的数码相机图像信号处理（ISP）流程包含如图 4-1 所示的若干个运算及处理模块<sup>[140, 141]</sup>，其中预处理模块通常包括非线性修正、坏点修正等操作，而后处理模块则包含图像增强、压缩、输出编码等操作。

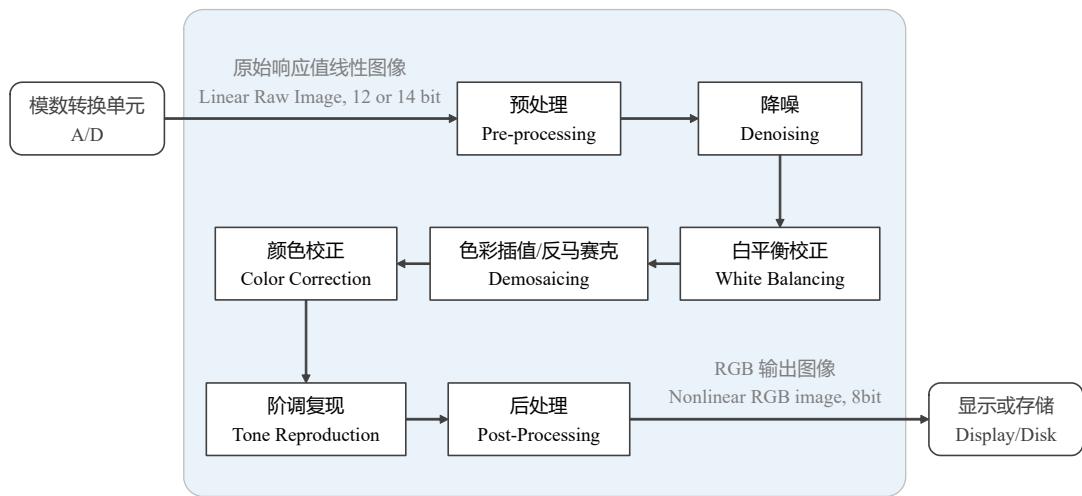


图 4-1 典型的数码相机图像信号处理流程

在图 4-1 所示的 ISP 流程中，色彩插值模块将通过空间插值的方式从单通道的传感器阵列图像（sensor array image）中获取相同分辨率的三通道图像（下文称之为 raw 图像）。该空间插值操作将导致待处理的数据量增加为原始数据量的三倍，因此在实际的片上（on-chip）ISP 流程中，为保证运算的实时性，往往会对全分辨率的 raw 图像进行下采样操作从而得到其低分辨率版本，并在该低分辨率图像上完成降噪、白平衡校正等模块所需参数的计算。

为确保输入图像中含有尽可能丰富的颜色信息，本论文并未使用下采样操作生成低分辨率图像，而是使用传感器彩色滤色片阵列下采样（sensor color filter array binning）的方式（参见图 3-3）从单通道 CFA 图像中恢复出三通道 raw 图像，并以此作为 ISP 流程的输入图像。本论文中所采用的 ISP 流程如图 4-2 所示，其中噪声修正模块表示使用第 3.2.2 小节中所介绍的系统误差修正方法对输入的 raw 图像进行必要的预处理，而后处理模块则表示将颜色校正后的 CIE1931 XYZ 图像进行必要的全局色调调节，并转换至目标 RGB 颜

色空间（如 sRGB 或 Adobe RGB）中，同时对其进行必要的输出编码，使之能够以图像文件的形式存储于硬盘或展示于显示器之中。

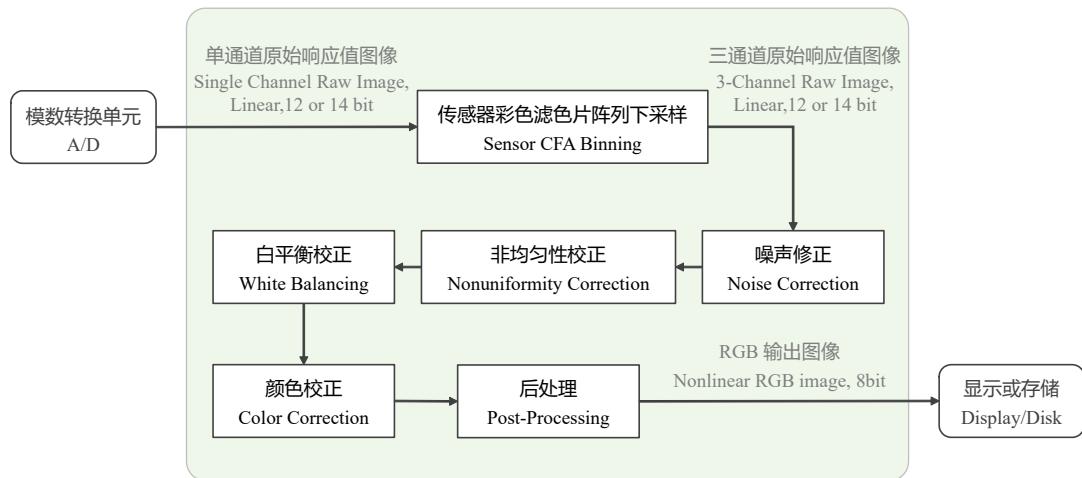


图 4-2 本论文所采用的图像信号处理流程

## 4.2 空间非均匀性校正

在第三章中我们简单地使用照度非均匀性调制函数  $U(i, j)$  对传感器平面上的空间非均匀性进行隐式的表示<sup>\*</sup>，本节中将对其成因及修正方法进行更加深入的探究。

得益于合理的光学系统设计，在大多数专业的单反或微单相机中，空间非均匀性并不显著。然而，对于移动终端中的相机而言，出于结构紧凑性的考虑，通常会将光学系统的出瞳与传感器平面间的距离设计得很小，导致传感器边缘区域对应了很大的主光线夹角<sup>[114]</sup>，因此此类成像系统的空间非均匀性较之专业相机也更加剧烈。为了对成像系统中的空间非均匀性进行更加直观的展示，本节将以移动终端中使用的 OmniVision OV8858 成像模组作为示例，其基本参数如表 4-1 所示，其中传感器分辨率一栏中括号中的数值表示经彩色滤色片阵列下采样后 RGB raw 图像的实际分辨率。

<sup>\*</sup> 事实上，“空间非均匀性”这一说法有可能令读者产生一定的歧义——对于具体的成像应用而言，最终图像中的空间非均匀性既有可能是由于成像系统自身的原因而导致的，也有可能是由于场景光照不均匀而导致的。显然，基于本论文的研究目标，我们仅希望对前者进行校正。针对成像系统自身的空间非均匀性，业内习惯使用“camera shading”或“lens shading”进行表述，然而为了避免文中出现“相机阴影”、“镜头阴影”这类更易引起误解的称谓，本论文中仍然采用了“空间非均匀性”这一说法。

表 4-1 空间非均匀性校正实验中所使用的测试相机相关参数

成像模组型号	OmniVision OV8858
传感器尺寸	4.735mm×3.517mm
传感器分辨率	4208×3120 (2104×1560×3)
彩色滤色片阵列格式	Bayer RGGB
实验中使用光圈数	f/2.0

#### 4.2.1 亮度非均匀性与颜色非均匀性

从图像处理的角度进行分析，相机系统的空间非均匀性可分为亮度非均匀性 (luminance shading) 与颜色非均匀性 (color shading) 两类，前者指的是图像三通道同时受到的整体强度调制，而后者则指的是不同位置处由于三通道相对强度发生变化而引起的颜色调制，如图 4-3 所示。



图 4-3 OV8858 成像模组的亮度非均匀性（左）与颜色非均匀性（右）

##### 4.2.1.1 亮度非均匀性

在摄影领域，图像边缘区域的响应值强度相较于中心区域存在一定程度的衰减这一现象通常被称为渐晕 (vignetting)。根据产生机理的不同，渐晕可分为自然渐晕与光学渐晕两类。

自然渐晕 (natural vignetting) 对应了式 (3.5) 中的四次余弦定理，本节中不再对其成因进行赘述。此类辐照度衰减是由光学成像系统的内在属性所决定的，因此对于任何相机系统而言都必然存在。需要注意的是，对于实际的光学系统而言，图 3-2 中所标注的主光线夹角  $\alpha$  在物方与像方往往并不相同，因此即使对于焦距相同的两支镜头 (对应相同的物方主光线夹角)，其自然渐晕的剧烈程度也可能存在差异。图 4-4 展示了不同的光学结构设计对于像方主光线夹角的影响。

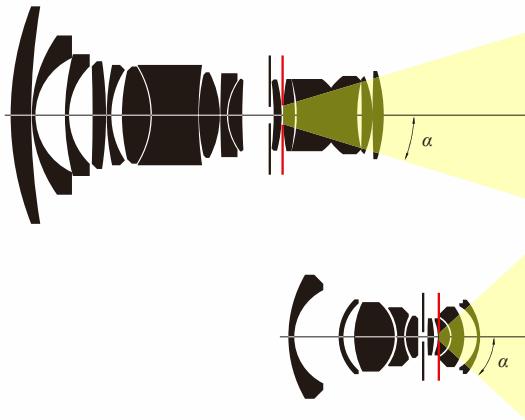


图 4-4 Carl Zeiss Distagon 21mm f/2.8 镜头（上）与 Carl Zeiss Biogon 21mm f/2.8 镜头（下）的光学结构示意图，其中黑色竖线表示孔径光阑，红色竖线表示出瞳。由于 Biogon 21mm f/2.8 具有更大的像方主光线夹角，其自然渐晕的程度要明显强于 Distagon 21mm f/2.8 © Paul van Walree<sup>[99]</sup>

为了缓解传感器边缘区域的照度衰减，部分传感器还会采用在每个感光单元彩色滤色片上方放置微透镜的方法用于汇聚边缘光线，从而增大位于边缘区域的感光单元的有效感光面积。对于大部分数码相机而言，其 raw 图像中亮度衰减的程度与四次余弦曲面并不严格吻合。

光学渐晕（optical vignetting）对应了在大光圈情况下部分斜入射的光线将被镜头外壁遮挡从而导致传感器边缘区域辐照度进一步衰减这一现象。由于镜头中的孔径光阑通常与最前端的镜片并不重合，且镜片的直径总是有限的，因此在大光圈的情况下，不同方向的入射光线实际上对应了不同的有效通光孔径。如图 4-5 所示，当入射光线相对于光轴的偏离角越大，则其对应的有效通光孔径就越小，因此传感器边缘区域的照度衰减也就越强烈。



图 4-5 Nikkor 50mm f/1.4D 镜头光圈全开时光轴方向（左）以及斜入射方向（右）上的有效通光孔径对比 © Discover Digital Photography<sup>[142]</sup>

对于大部分镜头而言，减小孔径光阑可以有效地缓解甚至消除此类渐晕。如图 4-6 所示，当使用 f/8.0 的光圈时，斜入射的光线不再受到镜头外壁的遮挡，此时有效的通光孔径完全取决于孔径光阑的大小。



图 4-6 Nikkor 50mm f/1.4D 镜头在 f/8.0 光圈下光轴方向（左）以及斜入射方向（右）上的有效通光孔径对比 © Discover Digital Photography [142]

此外，当镜头的遮光罩过长时，一部分视场边缘的光线将被遮光罩遮挡而无法进入相机系统，从而导致图像出现明显的暗角。此类现象通常被称为机械渐晕（mechanical vignetting）。通过取下遮光罩或者换用较短的遮光罩的方法即可完全消除机械渐晕。

#### 4.2.1.2 颜色非均匀性

在 ISP 流程中，除了需要对渐晕现象所导致的强度非均匀性进行补偿之外，还需要对 raw 图像不同位置处的颜色非均匀性进行校正。颜色非均匀性的主要成因在于不同波长的入射光线经滤色片与微透镜后到达光敏面的相对比例并不固定，而是随着像方主光线夹角的改变而改变<sup>[143]</sup>。因此，即使相机系统所拍摄的对象是一个光谱辐亮度固定不变的物体，随着该物体在传感器平面上的成像位置发生变化，其三通道原始响应值之间的相对强度也将发生变化，从而导致相机所记录的颜色产生变化。例如，对于某些相机，在主光线入射角较大时，其中波及短波段的入射光能量衰减较为剧烈，而长波段衰减则较为轻微，因此传感器边缘区域的 R 通道感光单元所接收的能量将略强于 G、B 通道所接收的能量（相对于传感器中央区域而言），此时即使对于一面均匀的白墙进行拍摄，所获得的 raw 图像也将呈现出中央偏青、周围偏红的颜色外貌。随着像方主光线入射角的增大，这种波长相关的能量强度相对差异也随之增大，因此对于移动终端相机等具有较大像方主光线夹角的成像系统，其颜色非均匀性也愈加显著。

式 (3.6) 假设了传感器平面上的非均匀性函数  $U(u, v)$  仅仅是关于位置坐标的函数<sup>\*</sup>，然而对于实际的相机系统而言，该调制因子还是关于入射能量光谱分布的函数<sup>[38]</sup>。一方面，由于不同颜色通道的感光单元对应了不同的光谱透过率函数  $t_k(\lambda), k \in \{r, g, b\}$ ，因此在三通道 raw 图像中，各颜色通道也将对应不同的非均匀性函数  $U_k(u, v), k \in \{r, g, b\}$ ；另一方

<sup>\*</sup> 为便于对样条曲面进行参数化表示，本节中使用 “ $u$ ”、“ $v$ ” 作为位置坐标符号。

面，当被摄物体自身的光谱辐亮度发生变化（例如照明光源发生了变化）时，各颜色通道的非均匀性函数也将发生变化。图 4-7 中展示了 OV8858 成像模组在不同光源下对均匀乳白玻璃拍摄得到的 raw 图像中  $x$  方向上的三通道响应值变化情况（该剖面线恰好经过光学中心）。

理论上，只有在已知成像系统光学传递函数的前提下才有可能通过光线追迹的方法对各通道间的非均匀性函数进行准确的标定。鉴于颜色非均匀性只在图像中含有大面积单色物体表面的情况下较为明显<sup>[144]</sup>，因此在实际的标定流程中，通常只选取若干种参考光源对均匀白场（flat field）进行标定。由于均匀白场中各点具有相同的光谱辐亮度，因此我们可将相机拍摄的白场 raw 图像  $\mathcal{I}_k^{\text{flat}}(u, v), k \in \{r, g, b\}$  作为相机系统在该光源下的非均匀性函数  $U_k(u, v), k \in \{r, g, b\}$  的近似。

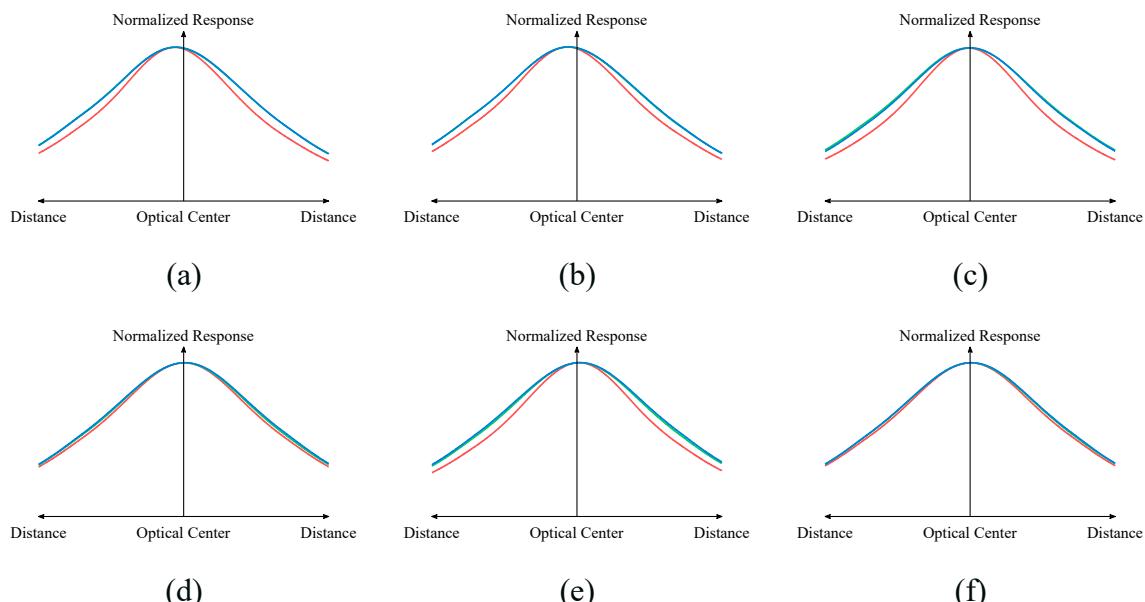


图 4-7 OV8858 成像模组在 (a) 4100K 卤素灯, (b) 7600K 卤素灯, (c) CIE A 模拟器, (d) CWF, (e) D65 模拟器, 以及 (f) TL84 光源下的白场图像各通道响应值随像素位置变化曲线（仅绘出  $x$  方向上的变化），其中(a)、(b)光源由 Image Engineering LE6-100 积分球配合不同色温的滤色片生成，(c)~(f)为 X-Rite Macbeth SpectraLight III 灯箱的内置光源

由上文可知，亮度非均匀性（渐晕）对图像各通道响应值的整体强度进行调制，而颜色非均匀性则对各通道之间的相对强度进行调制，因此在实际的标定实验中，只需要对三通道的非均匀性函数  $U_k(u, v), k \in \{r, g, b\}$  分别进行估计，即可同时完成亮度非均匀性与颜色非均匀性的标定。对于待校正的 raw 图像，只需要将其各颜色通道逐像素乘上估计出的

非均匀性函数的倒数（以下称之为倒数非均匀性函数） $\frac{1}{U_k(u,v)}$ ,  $k \in \{r,g,b\}$ ，即可实现空间非均匀性校正。

考虑到颜色非均匀性的光源相关性，在实际的实施过程中，通常会选取光源光谱功率分布具有代表性的若干种参考光源分别对三通道对应的非均匀性函数进行标定。在标定时，为了获得足够均匀的白场，需通过在光源上安置柔光罩或毛玻璃的方法对其做漫射处理，同时还需要在相机镜头前再放置一块乳白玻璃。此外，在对非均匀性函数进行标定时，为了获得较高的信噪比，应合理地设置曝光参数以确保白场图像在不饱和的前提下具有尽可能高的响应值。最后，由于光学系统的光圈大小将会对非均匀性函数 $U_k(u,v)$ 的强度与形状产生影响，因此对于光圈可变的成像系统，还需要针对不同光圈数分别进行标定。

#### 4.2.2 基于 B-样条曲面的空间非均匀性校正算法

为了对任意输入的 raw 图像进行空间非均匀性校正，需要已知该光源所对应的倒数非均匀性函数 $\frac{1}{U_k(u,v)}$ ,  $k \in \{r,g,b\}$ 。最简单的一种确定倒数非均匀性函数的方法是直接使用逐像素求倒数后的白场图像（以下称之为倒数白场图像） $\frac{1}{I_k^{\text{flat}}(u,v)}$ ,  $k \in \{r,g,b\}$ ，此时，对于同一光源下拍摄的任意输入图像，只需将其原始响应值逐像素地乘上 $\frac{1}{I_k^{\text{flat}}(u,v)}$ ，即可得到非均匀性校正后的图像。然而，一幅真实拍摄获得的 raw 图像中必然包含一定的噪声，同时，实验室环境下搭建的均匀白场也难以保证绝对的均匀性（例如乳白玻璃上存在污渍、毛刺等），因此若直接将白场图像作为空间非均匀性的校正依据，反而有可能引入一定程度的误差。在工程上，通常采用参数化的曲面方程对成像系统的空间非均匀性函数进行表示。这种参数化的函数表示方法既确保了非均匀性函数的空间平滑性，同时也大大减少了表征该非均匀性曲面所需要的数据量，因此允许我们对多种拍摄条件（例如不同的光源、不同的光圈数、不同的焦距，甚至不同的对焦距离等）分别进行标定。

部分研究认为 raw 图像中各通道的倒数非均匀性函数通常可以用一个径向对称的曲面进行描述<sup>[35, 145]</sup>，因此图像各位置处的响应值衰减程度将完全由其与光学中心的距离所决定。然而，这一假设仅适用于几何结构严格满足轴向对称的成像模组。对某些存在一定装配误差的模组而言，其非均匀性曲面并不规整，此时若使用径向对称的曲面对其进行拟合则很难保证足够的精度，甚至有可能在校正阶段引入一些明显的图像瑕疵。

为了确保非均匀性校正模型能够适用于不同的成像系统，本论文中采用 B-样条（B-

form spline) 曲面对倒数非均匀性函数  $\frac{1}{U_k(u,v)}$ ,  $k \in \{r, g, b\}$  进行拟合<sup>[146]</sup>。

#### 4.2.2.1 B-样条曲面拟合

B-样条曲面采用分块的方式对二维函数（即曲面）进行参数化表征，且保证了各曲面块连接处高阶导数的连续性，因此非常适合被用作参数模型对形状不规整的倒数非均匀性函数进行拟合。

一个二维 B-样条曲面  $B(u,v)$  可由一个  $(n_x+1) \times (n_y+1)$  控制点矩阵  $\mathbf{P}$  控制<sup>[147]</sup>：

$$B(u,v) = \sum_{x=0}^{n_x} \sum_{y=0}^{n_y} N_{x,d_x}(u) N_{y,d_y}(v) P(x,y) \quad (4.1)$$

式中  $P(x,y)$  表示矩阵  $\mathbf{P}$  中第  $y$  行  $x$  列的元素，且有  $0 \leq x \leq n_x$  以及  $0 \leq y \leq n_y$ ； $d_x$  与  $d_y$  表示样条曲面的阶数，且必须满足  $1 \leq d_x \leq n_x$  以及  $1 \leq d_y \leq n_y$ ，本节中我们令  $d_x = d_y = 3$ ； $N_{x,d_x}(u)$  与  $N_{y,d_y}(v)$  分别表示  $x$  与  $y$  方向上的 B-样条基函数 (B-spline basis functions)。在给定一组节点 (knot) 坐标  $\{t_i | 0 \leq i \leq n+d+1\}$  之后<sup>\*</sup>，基函数  $N_{i,j}(t)$  可通过递归计算的方式获得<sup>[147]</sup>：

$$N_{i,0}(t) = \begin{cases} 1, & t_i \leq t \leq t_{i+1} \\ 0, & \text{otherwise} \end{cases} \quad \text{for } 0 \leq i \leq n+d \quad (4.2)$$

以及

$$N_{i,j}(t) = \frac{t-t_i}{t_{i+j}-t_i} N_{i,j-1}(t) + \frac{t_{i+j+1}-t}{t_{i+j+1}-t_{i+1}} N_{i+1,j-1}(t) \quad (4.3)$$

for  $1 \leq j \leq d$  and  $0 \leq i \leq n+d-j$

对于倒数非均匀性函数拟合这一应用，节点坐标直接决定了曲面的分块规则。考虑到成像系统的空间非均匀性通常在图像的边缘区域较为剧烈，本论文采用双向非等距六分法对倒数白场图像  $\frac{1}{I^{\text{flat}}(u,v)}$  进行分块，如图 4-8 所示，此时  $x$ 、 $y$  方向的节点归一化坐标均为  $[0, 0, 0, .15, .3, .5, .7, .85, 1, 1, 1]$ 。对于一些空间非均匀性更加剧烈或更加不规整的成像系统，若有必要，也可采用更加密集的分块方式。

给定基函数、阶数以及  $k$  通道的倒数白场图像  $\frac{1}{I_k^{\text{flat}}(u,v)}$  后，我们使用最小二乘法对  $k$  通

<sup>\*</sup> 为方便起见，我们仅对某一方向上的基函数进行讨论，因此改用 “ $t$ ” 表示自变量。

道对应的 B-样条曲面控制点矩阵  $\mathbf{P}_k$  进行计算：

$$\hat{\mathbf{P}}_k = \arg \min_{u=1}^W \sum_{v=1}^H \left( \sum_{x=0}^{n_x} \sum_{y=0}^{n_y} N_{x, d_x}(u) N_{y, d_y}(v) P_k(x, y) - \frac{1}{\mathcal{I}_k^{\text{flat}}(u, v)} \right)^2, \quad k \in \{r, g, b\} \quad (4.4)$$

式中  $W$  与  $H$  分别表示倒数白场图像在  $x$  与  $y$  方向上的像素数。

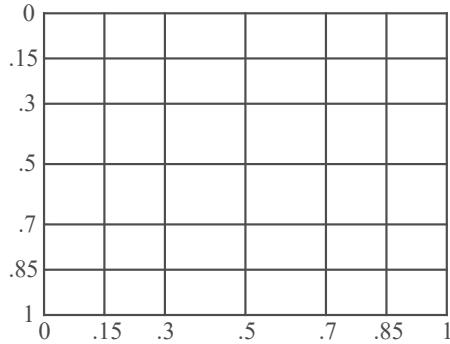


图 4-8 OV8858 成像模组使用 B-样条曲面拟合时所采用的图像分块规则

将式 (4.4) 中计算得到的  $\hat{\mathbf{P}}_k$  代入式 (4.1) 中，即可得到用于补偿 raw 图像中  $k$  通道空间非均匀性的 B-样条曲面  $\hat{B}_k(u, v)$ 。

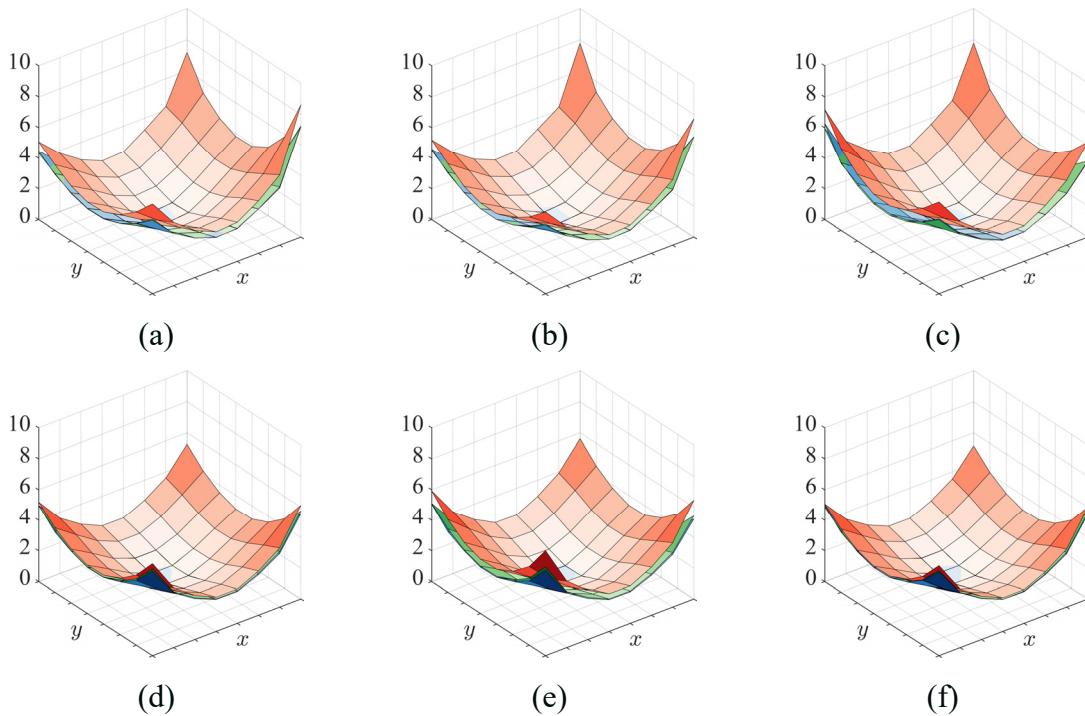


图 4-9 OV8858 成像模组在 (a) 4100K 卤素灯、(b) 7600K 卤素灯、(c) CIE A 模拟器、(d) CWF、(e) D65 模拟器以及 (f) TL84 光源下的各通道 B-样条补偿曲面参数可视化。对于以上六种标定光源下的白场图像，使用 B-样条曲面进行空间非均匀性校正后其边缘处的响应值衰减均在 95% 以上

当阶数  $d_x = d_y = 3$  且采用六分法对图像进行分块时, 有  $n_x = n_y = 7$ , 因此  $\mathbf{P}_k$  是一个  $8 \times 8$  矩阵。图 4-9 对不同光源下 B-样条补偿曲面对应的三通道控制点矩阵中的元素数值  $P_k(x, y)$  进行了可视化展示。

图 4-10 展示了六种标定光源的白场 raw 图像 G 通道在空间非均匀性校正前后的响应值分布情况。

#### 4.2.2.2 光源自适应的 B-样条曲面重构

显然, 在给定阶数以及节点坐标的情况下, 一个控制点矩阵  $\mathbf{P}$  唯一地确定一个 B-样条补偿曲面。正如图 4-7 与图 4-9 中所展示的那样, 同一个成像系统在不同光源下的空间非均匀性函数往往具有不同的形状, 因此用于重构补偿曲面的控制点矩阵也各不相同。为了使非均匀性校正模块能够根据不同的光源自适应地重构出补偿曲面, 我们需要获得场景光源与对应的控制点矩阵之间的对应关系。

为了将白平衡校正模块 (详见第 4.3 节) 所提供的光源估计信息用于空间非均匀性校正, 本论文使用相机 RGB 颜色空间中的 R、B 通道增益系数  $\mathbf{G} = [\mathcal{G}_r, \mathcal{G}_b]$  作为控制点矩阵的选取依据\*, 因此建立一个光源自适应的 B-样条曲面重构模型, 本质上就是建立起一组光源增益系数至控制点矩阵的映射  $f: \mathbf{G} \rightarrow \mathbf{P}$ 。

首先利用主成分分析 (PCA) 对控制点矩阵进行降维。对于某个标定光源的  $k$  通道控制点矩阵  $\mathbf{P}_k, k \in \{r, g, b\}$ , 我们将其各元素按列展开并转置, 因此可得到一个长度为  $(n+1)^2$  的控制点行向量 (假设  $n_x = n_y = n$ )。将  $M$  个标定光源的控制点行向量沿纵向进行组合, 可得到用于主成分分析的  $M \times (n+1)^2$  控制点组合矩阵  $\mathbf{Q}_k$ :

$$\mathbf{Q}_k = \begin{bmatrix} P_k^{ill_1}(1,1) & \cdots & P_k^{ill_1}(n+1,1) & \cdots & P_k^{ill_1}(1,n+1) & \cdots & P_k^{ill_1}(n+1,n+1) \\ \vdots & & \vdots & & \vdots & & \vdots \\ P_k^{ill_M}(1,1) & \cdots & P_k^{ill_M}(n+1,1) & \cdots & P_k^{ill_M}(1,n+1) & \cdots & P_k^{ill_M}(n+1,n+1) \end{bmatrix}, k \in \{r, g, b\} \quad (4.5)$$

利用奇异值分解 (SVD) 对  $\mathbf{Q}_k$  进行主成分分析, 可得到一个  $M \times M$  的权重矩阵  $\mathbf{W}_k$  以及一个  $(n+1)^2 \times M$  的基向量矩阵  $\mathbf{X}_k$ , 且满足

$$\mathbf{Q}_k = \mathbf{W}_k \cdot \mathbf{X}_k^T, \quad k \in \{r, g, b\} \quad (4.6)$$

\* 在光源颜色估计任务中, 研究者们通常更关心光源色三通道间的相对强度关系, 因此在计算三通道的增益系数时往往将 G 通道的增益系数固定为 1, 即  $\mathcal{G}_g \equiv 1$ , 此时使用 R、B 通道的一对增益系数  $[\mathcal{G}_r, \mathcal{G}_b]$  即可对光源颜色进行表征。

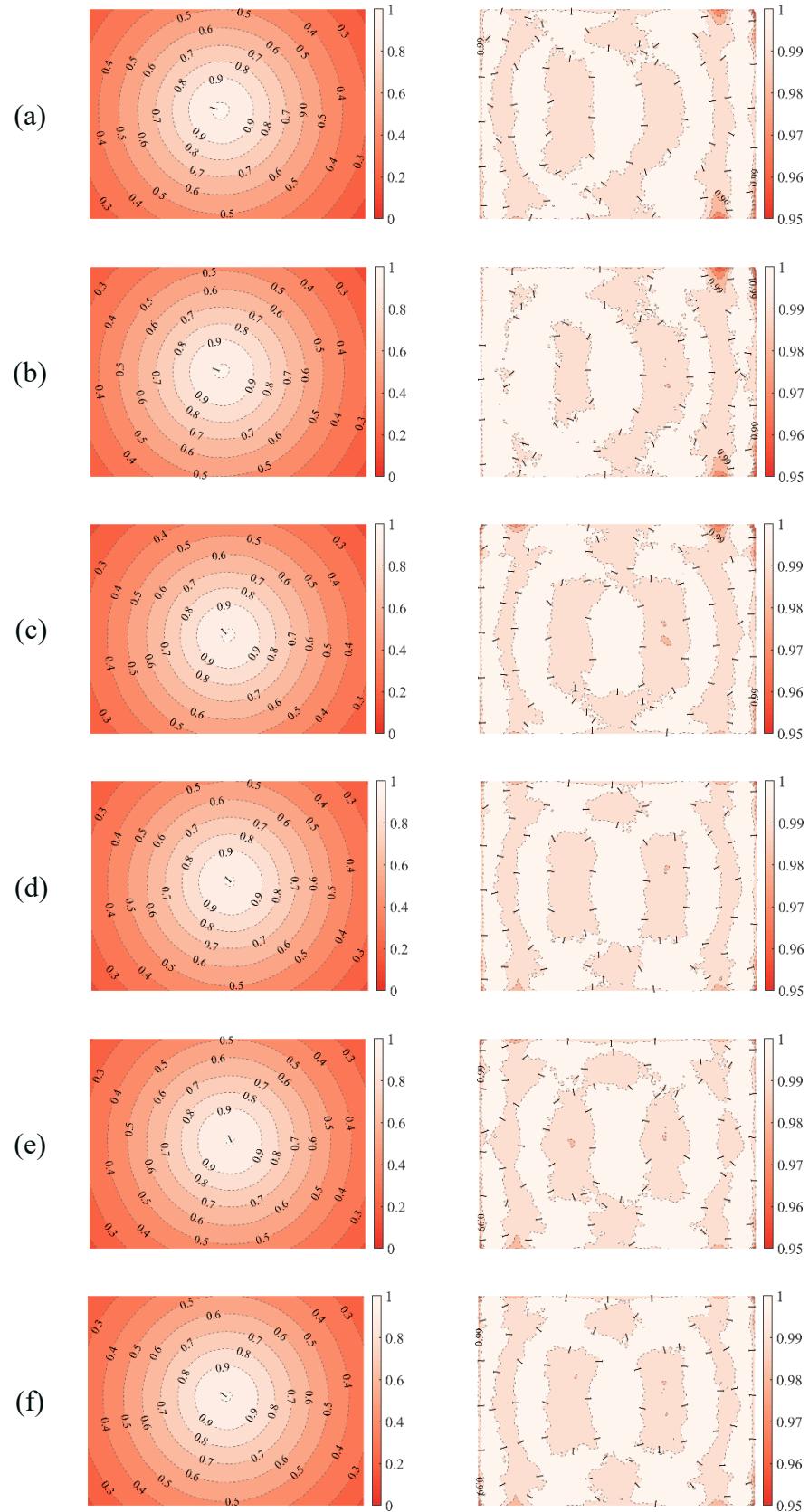


图 4-10 OV8858 成像模组在 (a) 4100K 卤素灯、(b) 7600K 卤素灯、(c) CIE A 模拟器、(d) CWF、(e) D65 模拟器以及 (f) TL84 光源下非均匀性校正前 (左) 后 (右) G 通道响应值分布对比

式中  $\mathbf{X}_k$  中每一列代表一个基向量（主成分），且各个基向量对于恢复  $\mathbf{Q}_k$  所起的重要程度随着列数的增加而依次递减； $\mathbf{W}_k$  的每一行则代表着一个标定光源对应的  $M$  个基向量权重系数：

$$\mathbf{W}_k = \begin{bmatrix} W_{k,1}^{ill_1} & W_{k,2}^{ill_1} & \cdots & W_{k,M}^{ill_1} \\ \vdots & \vdots & & \vdots \\ W_{k,1}^{ill_M} & W_{k,2}^{ill_M} & \cdots & W_{k,M}^{ill_M} \end{bmatrix}, \quad k \in \{r, g, b\} \quad (4.7)$$

当标定光源的数量  $M$  较大时，通常只需要保留具有最大特征值的前  $N$  个基向量以及对应的权重系数，即可以足够小的误差恢复出控制点组合矩阵  $\mathbf{Q}_k$ ，此时  $\mathbf{W}_k$  成为  $M \times N$  矩阵，而  $\mathbf{X}_k$  成为  $(n+1)^2 \times N$  矩阵。由于本节中使用的标定光源数并不算太多 ( $M=6$ )，因此在执行完主成分分析后仍然选择保留所有的基向量以及权重系数，即令  $N=M$ 。

显然，在给定基向量矩阵  $\mathbf{X}_k$  的情况下，一个  $k$  通道的  $(n+1) \times (n+1)$  控制点矩阵  $\mathbf{P}_k^{ill_x}$  将由一组权重系数  $[W_{k,1}^{ill_x}, \dots, W_{k,M}^{ill_x}]$  唯一决定，因此用于重构一个 B-样条曲面所需的参数数量从  $(n+1)^2$  减少至  $M$ 。换句话说，一旦确定了 B-样条曲面的阶数、节点坐标以及标定光源的一组主成分（即基向量矩阵），只需  $M$  个参数便可重构出一个空间非均匀性补偿曲面。

在获得基向量与相应的权重系数之后，使用非齐次多项式模型作为增益系数  $\mathbf{G} = [\mathcal{G}_r, \mathcal{G}_b]$  至 PCA 权重系数的映射模型：

$$\begin{cases} W_{k,1} = a_{k,10} + a_{k,11}\mathcal{G}_r + a_{k,12}\mathcal{G}_b + a_{k,13}\mathcal{G}_r^2 + a_{k,14}\mathcal{G}_b^2 + a_{k,15}\mathcal{G}_r\mathcal{G}_b \\ \vdots \\ W_{k,M} = a_{k,M0} + a_{k,M1}\mathcal{G}_r + a_{k,M2}\mathcal{G}_b + a_{k,M3}\mathcal{G}_r^2 + a_{k,M4}\mathcal{G}_b^2 + a_{k,M5}\mathcal{G}_r\mathcal{G}_b \end{cases}, \quad k \in \{r, g, b\} \quad (4.8)$$

式中各系数可利用最小二乘法求得：

$$\mathbf{a}_k = \begin{bmatrix} a_{k,10} & \cdots & a_{k,M0} \\ a_{k,11} & \cdots & a_{k,M1} \\ \vdots & & \vdots \\ a_{k,15} & \cdots & a_{k,M5} \end{bmatrix} = (\mathbf{G}_{\text{poly}})^+ \cdot \mathbf{W}_k, \quad k \in \{r, g, b\} \quad (4.9)$$

其中  $(\cdot)^+$  表示计算伪逆矩阵 (Moore–Penrose pseudoinverse matrix)，且有  $\mathbf{X}^+ = \mathbf{X}^\top(\mathbf{X}\mathbf{X}^\top)^{-1}$ ；

$\mathbf{G}_{\text{poly}}$  表示由  $M$  个标定光源的增益系数多项式构成的  $M \times 6$  矩阵：

$$\mathbf{G}_{\text{poly}} = \begin{bmatrix} 1 & \mathcal{G}_r^{ill_1} & \mathcal{G}_b^{ill_1} & (\mathcal{G}_r^{ill_1})^2 & (\mathcal{G}_b^{ill_1})^2 & \mathcal{G}_r^{ill_1}\mathcal{G}_b^{ill_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \mathcal{G}_r^{ill_M} & \mathcal{G}_b^{ill_M} & (\mathcal{G}_r^{ill_M})^2 & (\mathcal{G}_b^{ill_M})^2 & \mathcal{G}_r^{ill_M}\mathcal{G}_b^{ill_M} \end{bmatrix} \quad (4.10)$$

### 4.2.2.3 空间非均匀性校正流程

在完成以上标定过程后，对于任意给定的输入 raw 图像  $\mathcal{I}$  以及对应的白平衡增益系数  $\mathbf{G} = [\mathcal{G}_r, \mathcal{G}_b]$ （由自动白平衡校正模块进行估计），我们采取以下流程对其空间非均匀性进行校正：

- 1) 利用式 (4.8) 计算出各通道的基向量对应的权重系数

$$\mathbf{W}_k^{\mathcal{I}} = [W_{k,1}, \dots, W_{k,M}], \quad k \in \{r, g, b\} \quad (4.11)$$

式中  $\mathbf{W}_k^{\mathcal{I}}$  是一个  $M \times 1$  的行向量。

- 2) 利用

$$\mathbf{Q}_k^{\mathcal{I}} = \mathbf{W}_k^{\mathcal{I}} \cdot \mathbf{X}_k^{\top}, \quad k \in \{r, g, b\} \quad (4.12)$$

以及

$$\mathbf{P}_k^{\mathcal{I}} = \text{reshape}\left[\left(\mathbf{Q}_k^{\mathcal{I}}\right)^{\top}\right], \quad k \in \{r, g, b\} \quad (4.13)$$

计算出该图像中的估计光源对应的各通道的控制点矩阵  $\mathbf{P}_k^{\mathcal{I}}$ ，其中  $\text{reshape}(\cdot)$  表示将长度为  $(n+1)^2$  的列向量重新排列为一个  $(n+1) \times (n+1)$  的矩阵。

- 3) 利用式 (4.1) 重构出用于空间非均匀性补偿的 B-样条曲面  $B_k(u, v), k \in \{r, g, b\}$ 。
- 4) 将  $B(u, v)$  与输入图像  $\mathcal{I}$  进行逐通道、逐元素相乘，得到空间非均匀性校正后的输出图像。

图 4-11 展示了使用上文所述的方法对测试图像进行空间非均匀性校正前后的对比结果。



图 4-11 不同测试场景下 OV8858 成像模组拍摄的 raw 图像在空间非均匀性校正前（左）后（右）的结果对比

### 4.3 颜色恒常性与自动白平衡校正

从辐射度学的角度进行分析，当同一物体位于不同光源下时，其二次反射出的光谱辐亮度也将具有不同的分布，因而对应了不同的色度三刺激值。然而，人眼视觉系统在观察物体时，往往能够迅速而准确地排除照明体色品对于该物体颜色外貌的干扰，从而获得该物体在参考光源下的感知颜色外貌。这种视觉特性通常被称为色适应 (chromatic adaptation) 或颜色恒常性 (color constancy)。

一些研究者认为视觉系统中的颜色恒常性是进化与适应的产物<sup>[148-150]</sup>，即人类在长期进化的过程中出于在各种野外复杂环境中生存的需要，而逐渐演化出这种光源适应的能力<sup>[151]</sup>，其中大多数观察者选择日光这一自然光源作为“参考光源”便是该理论的有力证明<sup>[71]</sup>。

[152]。然而，颜色恒常性的生物学机理至今仍然没有得到彻底、全面的解释[153]。在已有一些理论分析中，记忆色理论受到了大部分研究者的认同[154-156]，即，颜色恒常性是人眼视觉系统与大脑共同协作的产物，当观察者在判断某一物体颜色时，通常会以其记忆中对于此类物体的颜色（甚至是光谱反射比[155]）的感知作为参照，从而推断出所观察的物体对应的“真实颜色”。例如，在15000K以上色温的光源下观察一根香蕉，即使此时香蕉表面反射光谱中的短波能量远大于中长波能量，但是只要观察者识别出了这是一根香蕉，他/她仍然能够做出“香蕉皮是黄色的”这一判断。

颜色恒常性在数字成像领域同样也扮演着重要角色。对于摄影应用，图像中物体的颜色外貌与用户真实感知的颜色外貌之间是否一致，将显著影响用户对于图像质量的评价[157]；对于计算机视觉应用，能否准确地排除光源干扰、恢复出物体真实颜色，将直接关系到图像识别、分类等算法的性能。然而，数码相机在硬件层面并不具备颜色恒常性这一特性，其记录的物体颜色将严格遵从式(3.15)中的积分模型，因此在与计算摄影学相关的应用中，必须从软件层面对颜色恒常性进行模拟，从而使最终的输出图像具有与人眼视觉系统感知尽可能一致的颜色外貌。

广义上的数字图像颜色恒常性指的是将任意光源下捕获的图像校正至该场景在参考光源下所呈现出的颜色外貌特征。在ISP流程中，这一过程通常需要被分解为两个模块分别进行处理：第一个模块仅对图像中的中性色（neutral color）进行调节，而不追求其它颜色（chromatic color）的准确还原；第二个模块则在不改变前一模块已校正好的中性色的前提下对其余颜色进行调节。由于第一个模块输出图像中的中性色物体（例如白墙、灰卡等）已经具有了正确的颜色，因此这一过程也常被称为白平衡校正（white balancing 或 white balance correction）；对应地，第二个模块通常被称为颜色校正（color correction）。关于颜色校正的理论基础以及实现方法将在第4.4节中进行讨论。

在图像处理领域，颜色恒常性计算（computational color constancy）通常仅狭义地指对输入图像进行白平衡校正。

为了将未知光源下中性色物体的响应值 $D_{n,k}^U, k \in \{r, g, b\}$ 校正至参考光源下的响应值 $D_{n,k}^C, k \in \{r, g, b\}$ <sup>\*</sup>，需将其三通道各自乘上一个增益系数 $G_k, k \in \{r, g, b\}$ ：

---

<sup>\*</sup> 本节中，下标“n”表示中性色（neutral）物体，上标“U”表示未知（unknown）光源，“C”表示参考（canonical）光源。

$$\begin{cases} \mathcal{G}_r = D_{n,r}^C / D_{n,r}^U \\ \mathcal{G}_g = D_{n,g}^C / D_{n,g}^U \\ \mathcal{G}_b = D_{n,b}^C / D_{n,b}^U \end{cases} \quad (4.14)$$

由于理想的中性色物体的光谱反射比在各波长处恒为一常数，因此根据式 (3.60)，其线性原始响应值等价于光源的光谱功率分布函数与相机光谱灵敏度的积分乘上一常数，即

$$\begin{cases} D_{n,r}^U = c_r \int_{\lambda} E^U(\lambda) s_r(\lambda) d\lambda \\ D_{n,g}^U = c_g \int_{\lambda} E^U(\lambda) s_g(\lambda) d\lambda \\ D_{n,b}^U = c_b \int_{\lambda} E^U(\lambda) s_b(\lambda) d\lambda \end{cases} \quad (4.15)$$

式中  $E^U(\lambda)$  代表未知光源  $U$  的相对光谱功率分布函数\*。

为了避免乘上一个小于 1 的增益系数而导致图像中饱和区域出现偏色，在实际操作的过程中，通常需要将三通道白平衡增益系数同时乘上一个常数，以确保其中最小的一个系数恰好等于 1。以常见的 Bayer 阵列传感器为例，其 G 通道感光单元的灵敏度通常要大于 R、B 通道，因此  $\mathcal{G}_g$  通常小于  $\mathcal{G}_r$ 、 $\mathcal{G}_b$ ，此时有

$$\begin{cases} \mathcal{G}_r \leftarrow \mathcal{G}_r / \mathcal{G}_g \\ \mathcal{G}_g = 1 \\ \mathcal{G}_b \leftarrow \mathcal{G}_b / \mathcal{G}_g \end{cases} \quad (4.16)$$

在下文中，若无特别说明，则默认 G 通道的白平衡增益系数为 1，而仅使用  $\mathcal{G}_r$ 、 $\mathcal{G}_b$  对输入图像进行白平衡校正。由式 (4.15) 可知，在对 G 通道的白平衡增益系数进行归一化的情况下，计算光源 X 下中性色物体的响应值等价于计算光源 X 自身的响应值，因此在下文中，我们将使用“光源 X 的响应值”或“光源 X 的色度”的表达方式来代替“光源 X 下中性色物体的响应值”或“光源 X 下中性色物体的色度”的说法，同时，光源自身的响应值也就等价于带有“n”下标的中性色物体的响应值。

显然，根据式 (4.14)，在参考光源响应值  $D_{n,k}^C$  已知的情况下，对白平衡增益系数  $\mathcal{G}_r$ 、 $\mathcal{G}_b$  进行估计等价于对未知光源的响应值（或色度） $D_{n,k}^U$  进行估计，因此在一些文献中，自动白平衡校正问题也被称作光源颜色估计（illuminant color estimation）问题<sup>[11, 57, 158, 159]</sup>。

\* 在与颜色恒常性相关的应用中，研究者通常仅关心光谱功率分布函数的相对形状而非其幅值大小，因此本节中我们使用相对光谱功率分布函数作为光源的表征，并使用符号  $c$  来表示式中所需的所有常数项因子。

由于实际场景中的光源千差万别，由 $(D_{n,r}^U, D_{n,g}^U, D_{n,b}^U)$ 所构成的样本空间可近似视为连续空间，因此自动白平衡问题实质上是一个回归问题：

$$f: \mathcal{I} \rightarrow D_{n,r}^U, D_{n,g}^U, D_{n,b}^U \quad (4.17)$$

对于任意一幅 raw 图像，其各个像素的原始响应值是由光源（光谱功率分布函数）与被摄物体（光谱反射比函数）共同决定的。在没有任何约束的情况下，从输入图像  $\mathcal{I}$  中对光源颜色或白平衡增益系数进行估计是一个病态（ill-posed）问题，即有无穷多种光源与物体的组合能够满足当前图像的响应值分布。换言之，在不做出任何先验假设的前提下，永远无法判断一幅黄色的图像是由于黄光投射在白墙上产生的，还是由于白光投射在黄墙上产生的。这种本质上的不确定性也正是颜色恒常性计算问题的难点所在——永远不存在一种能够完美应对所用场景的自动白平衡算法。

#### 4.3.1 结合色域映射与中性色像素统计的自动白平衡算法（GMap-NPStat）

对于日常拍摄应用而言，大部分场景中总是或多或少含有中性色的物体表面<sup>[42, 49]</sup>。对此类场景，一旦能够确定中性色物体表面所对应的像素位置，便可利用式 (4.14) 对白平衡增益系数进行计算。此类基于中性色物体表面颜色信息的自动白平衡算法，本质上是对输入图像中的中性色像素进行搜索、筛选与统计。一种典型的中性色像素选取策略是，遍历输入图像中的每个像素，判断其二维色品是否落在一个预先设定好的中性色区域中，若是，则认为该像素属于中性色像素，并将其收集至候选中性色像素集合中，若否，则舍弃该像素；待遍历过程结束后，使用灰度世界算法<sup>[41]</sup>对候选中性色像素集合中的所有像素进行等权统计，从而得到该图像对应的白平衡增益系数  $G_r$ 、 $G_b$ 。

这种基于像素统计的自动白平衡算法简单、高效，然而其难点在于如何对中性色区域进行预先的设定：若中性色区域设定得过小，则难以保证输入图像中有足够数量的像素能够落入该区域内，从而影响了统计结果的可靠性；若中性色区域设定得过大，则势必将许多“伪中性色”像素囊括至候选中性色像素集合中，从而使统计结果出现偏差。

本节中，针对这一问题，首先设定一个足够大的、能够覆盖各种常见光源色品坐标的中性色区域，并利用色域映射（gamut mapping）算法<sup>[160]</sup>，根据当前输入图像自适应地调整该中性色区域的边界，在不遗漏真实的中性色像素的前提下尽可能缩小有效的中性色区域面积，以确保统计结果的准确性。我们将该自动白平衡算法命名为“结合色域映射与中性色像素统计的自动白平衡算法（Gamut Mapping-based Neutral Pixels Statistics）”，下文中

简称为 GMap-NPStat 算法。

#### 4.3.1.1 正交色度平面

改变场景中光源的强度或调整相机的曝光参数均有可能导致 raw 图像的整体亮度发生变化，因此在对中性色像素进行判断时，通常需要排除亮度信息的影响。类似于将 CIE1931 XYZ 三刺激值映射至 CIE1931 xy 色品平面，本小节使用一个设备相关的二维平面对 raw 图像中的色度信息进行表示。

为了避免中性色区域边界的选取策略过于复杂（例如为了在 CIE1931 xy 色品平面内确定一个包含黑体轨迹的“白点”区域，需要指定一个十分复杂的非凸多边形），我们希望该色度平面具有以下两个性质：

- 1) 当光源的相关色温发生变化时，该平面内的色度坐标仅沿水平方向发生变化；
- 2) 当光源的相关色温不变而色品坐标在等色温线上变化时，该平面内的色度坐标仅沿竖直方向发生变化。

当新的色度平面满足以上两点性质时，黑体轨迹与等色温线在该平面内近乎于垂直分布，因此我们将该平面称为正交色度平面（orthogonal chromatic plane, OCP），并使用符号  $X_{orth}$  与  $Y_{orth}$  对其坐标进行表示。

我们使用如下映射关系实现相机 RGB 原始响应值  $[D_r, D_g, D_b]$  至正交色度平面坐标  $[X_{orth}, Y_{orth}]$  的变换：

$$\begin{aligned} \begin{bmatrix} X_{orth} \\ Y_{orth} \end{bmatrix} &= \mathbf{S} \cdot \mathbf{R} \cdot \begin{bmatrix} X - X_0 \\ Y - Y_0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \sigma \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} \log \left( \frac{D_r}{D_r^C} \frac{w_r D_r^C + w_g D_g^C + w_b D_b^C}{w_r D_r + w_g D_g + w_b D_b} \right) \\ \log \left( \frac{D_b}{D_b^C} \frac{w_r D_r^C + w_g D_g^C + w_b D_b^C}{w_r D_r + w_g D_g + w_b D_b} \right) \end{bmatrix} \end{aligned} \quad (4.18)$$

式中  $[D_r^C, D_g^C, D_b^C]$  表示参考光源对应的相机 RGB 响应值<sup>\*</sup>， $w_k, k \in \{r, g, b\}$  表示各通道的权重系数， $\theta$  表示坐标旋转角度， $\sigma$  表示对正交色度平面中水平方向坐标进行调节的“修剪系数”， $\mathbf{R}$  与  $\mathbf{S}$  称为变换过程中的“旋转矩阵”与“修剪矩阵”。

本论文基于原始响应值预测模型对变换表达式 (4.18) 中的参数进行计算，其具体原

<sup>\*</sup> 本节中，在不引起歧义的情况下，对于光源自身的响应值我们默认省略下标 “n”。

理及实现过程请参阅第 5.1.1 小节。

图 4-12 展示了若干日光光源以及 LED 光源在 Nikon D3x 相机正交色度平面上的坐标分布，其中日光光源由相关色温由 3000K 至 12000K 的一系列重组日光光谱构成，其坐标在正交色度平面内大致沿水平方向分布；LED 光源由光谱可调的多通道 LED 平台<sup>[122]</sup>生成，其相关色温均为 6500K，但 CIE1931 xy 色品坐标在  $-0.02 \leq Duv \leq 0.02$ <sup>[161]</sup> 等色温线上均匀分布，因此其坐标在正交色度平面内大致沿竖直方向分布。

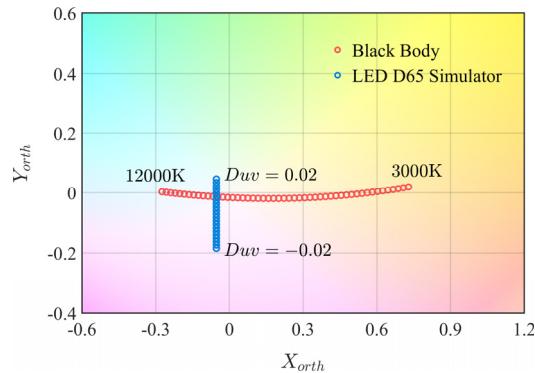


图 4-12 3000K 至 12000K 的重组日光光源以及  $Duv$  不同的多通道 LED 光源在 Nikon D3x 正交色度平面上的坐标分布

虽然下文中即将介绍的自动白平衡算法将在正交色度平面内进行中性色像素统计，但算法最终输出的增益系数  $G_r$ 、 $G_b$  仍然应对应于相机 RGB 颜色空间，因此，还应当考虑如何将正交色度平面坐标重新映射回相机 RGB 响应值。根据式 (4.18) 可知，在相机 RGB 颜色空间至正交色度平面的变换过程中响应值的亮度信息将被移除，因此理论上不可能由一组  $[X_{orth}, Y_{orth}]$  坐标唯一确定一组相机 RGB 响应值。所幸的是，在处理白平衡校正这一问题时，仅需要考虑光源 RGB 响应值三通道间的相对关系，因此不妨将 G 通道值设定为一常数，这样便可通过对式 (4.18) 求逆过程从而确定正交色度平面坐标至相机 RGB 响应值的逆变换关系。

具体地，在计算正交色度平面至相机 RGB 颜色空间的逆变换时，令  $D_g \equiv 1$ ，故有

$$\begin{bmatrix} D_r \\ D_b \end{bmatrix} = \begin{bmatrix} (e^X - 1)w_r & e^X w_b \\ e^Y w_r & (e^Y - 1)w_b \end{bmatrix}^{-1} \cdot \begin{bmatrix} -e^X w_g \\ -e^Y w_g \end{bmatrix} \quad (4.19)$$

式中， $X$  与  $Y$  满足

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \mathbf{R}^{-1} \cdot \mathbf{S}^{-1} \cdot \begin{bmatrix} X_{orth} \\ Y_{orth} \end{bmatrix} + \begin{bmatrix} \log \frac{w_r D_r^C}{w_r D_r^C + w_g D_g^C + w_b D_b^C} \\ \log \frac{w_b D_b^C}{w_r D_r^C + w_g D_g^C + w_b D_b^C} \end{bmatrix} \quad (4.20)$$

#### 4.3.1.2 基于二维直方图的中性色像素统计算法 (NPStat)

通过对正交色度平面进行二维网格划分并统计落入各个格点的像素数，能够得到输入图像  $\mathcal{I}$  在正交色度平面中的直方图分布。然而，这种对图像中的所有像素不加区分地进行累加的直方图生成方法并未考虑到像素的亮度信息在光源估计问题中所起的作用。基于图像中亮度较高的像素往往能够更加准确地反映光源的色度信息<sup>[162]</sup>这一假设，我们使用每个像素的“亮度”  $L = w_r D_r + w_g D_g + w_b D_b$  作为直方柱高度累加的权重生成最终的亮度加权直方图 (lightness weighted histogram)  $\mathbf{H}$ ，即

$$\begin{aligned} H(a, b) &= \sum_i^w \sum_j^h B(i, j) L^\gamma(i, j) \\ &= \sum_i^w \sum_j^h B(i, j) [w_r D_r(i, j) + w_g D_g(i, j) + w_b D_b(i, j)]^\gamma \end{aligned} \quad (4.21)$$

式中  $H(a, b)$  表示直方图  $\mathbf{H}$  中单元格  $(a, b)$  处的直方柱高度， $w$  与  $h$  分别表示图像  $\mathcal{I}$  的宽与高， $\gamma$  表示对亮度权重进行指数调制的因子（本论文中  $\gamma = 3$ ）， $B$  表示一个用于判断当前像素是否落在单元格  $(a, b)$  内的布尔变量：

$$B(i, j) = \begin{cases} 1 & \text{if } X_a^- \leq X_{orth}(i, j) < X_a^+ \text{ and } Y_b^- \leq Y_{orth}(i, j) < Y_b^+ \\ 0 & \text{otherwise} \end{cases} \quad (4.22)$$

式中  $X_a^+$  与  $X_a^-$  分别表示位于第  $a$  行的单元格在正交色度平面水平方向上有效统计区间的上、下限， $Y_b^+$  与  $Y_b^-$  同理。

图 4-13 比较了对于同一幅测试图像常规直方图与亮度加权直方图的区别，从中可以观察到，相比于常规的直方图生成方法，式 (4.21) 中的方法削弱了某些亮度较低的像素（该场景中主要为部分阴影处的深绿色植物）对直方图所起的贡献，并提升了某些亮度较高的像素（该场景中主要为背景中浅色的石墙）所占的权重，从而降低了某些伪中性色像素以及图像暗部的彩色噪声对光源颜色估计所造成的干扰。

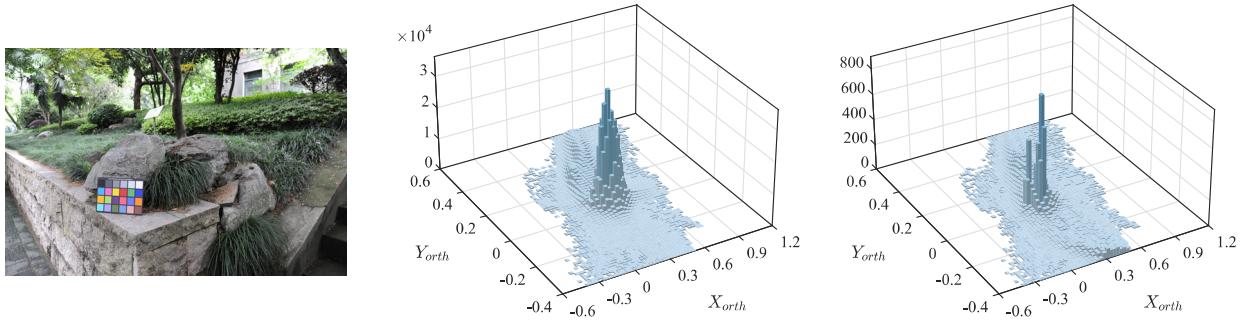


图 4-13 正交色度平面中测试图像（左）对应的常规直方图（中）与亮度加权直方图  $\mathbf{H}$ （右）

在得到输入图像对应的亮度加权直方图  $\mathbf{H}$  后，可筛选出其中位于正交色度平面原点（即参考光源 D65 对应的  $[X_{orth}, Y_{orth}]$  坐标）附近的像素作为候选中性色像素。具体地，我们利用一个用户预先设定的“候选中性色区域（candidate neutral region）”  $\mathcal{T}_n$  生成一个与亮度加权直方图  $\mathbf{H}$  具有相同尺寸的候选中性色区域掩模（candidate neutral region mask） $\mathbf{M}_n$ ，并将其与  $\mathbf{H}$  求交，从而将  $\mathbf{H}$  中所有位于候选中性色区域之外的直方柱置零，仅保留该区域内的统计数据：

$$\mathbf{H}_n = \mathbf{H} \circ \mathbf{M}_n \quad (4.23)$$

式中符号“ $\circ$ ”表示对两个矩阵进行逐元素相乘（Hadamard 乘积）。 $\mathbf{M}_n$  中各元素的数值满足以下条件：

$$M_n(a, b) = \begin{cases} 1 & \text{if } \left( \frac{X_a^- + X_a^+}{2}, \frac{Y_b^- + Y_b^+}{2} \right) \in \mathcal{T}_n \\ 0 & \text{otherwise} \end{cases} \quad (4.24)$$

我们将求交后的直方图  $\mathbf{H}_n$  称为候选中性色像素直方图（candidate neutral histogram）。

图 4-14 展示了一种可能的候选中性色区域  $\mathcal{T}_n$  以及最终获得的候选中性色像素直方图。

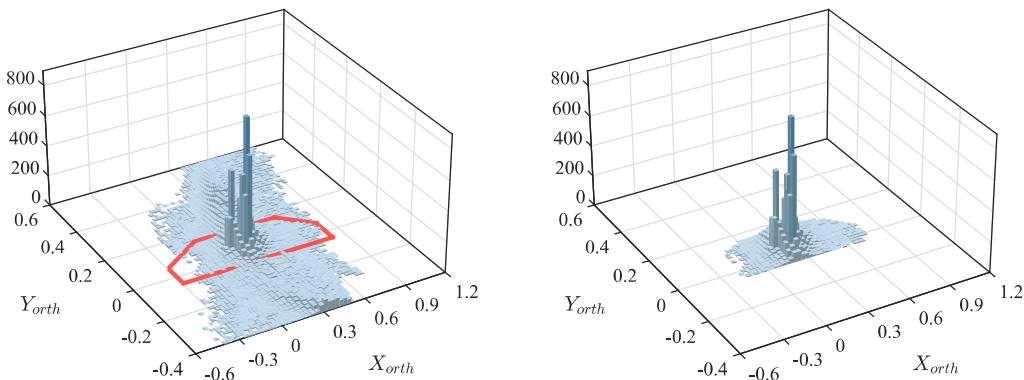


图 4-14 左：预先设定的候选中性色区域  $\mathcal{T}_n$ （红色实线）；右：求交后的候选中性色像素直方图  $\mathbf{H}_n$

本论文基于原始响应值预测模型对该候选中性色区域  $T_n$  进行确定，其具体原理及实现过程请参阅第 5.1.2 小节。

在得到候选中性色像素直方图  $\mathbf{H}_n$  后，将其重心坐标作为未知光源  $U$  在正交色度平面中的坐标估计：

$$\begin{cases} \hat{X}_{orth}^U = \frac{X^+ - X^-}{M_X} \hat{a} + X^- \\ \hat{Y}_{orth}^U = \frac{Y^+ - Y^-}{M_Y} \hat{b} + Y^- \end{cases} \quad (4.25)$$

式中

$$\hat{a} = \frac{\sum_{b=1}^{M_Y} \sum_{a=1}^{M_X} a H_n(a, b)}{\sum_{b=1}^{M_Y} \sum_{a=1}^{M_X} H_n(a, b)}, \quad \hat{b} = \frac{\sum_{b=1}^{M_Y} \sum_{a=1}^{M_X} b H_n(a, b)}{\sum_{b=1}^{M_Y} \sum_{a=1}^{M_X} H_n(a, b)} \quad (4.26)$$

分别表示  $\mathbf{H}_n$  的重心在直方图二维网格坐标系中水平与竖直方向上的坐标， $M_X$  与  $M_Y$  分别表示水平与竖直方向上的网格数， $X^+$  与  $X^-$  分别表示直方图的统计区间在正交色度平面水平方向上的上、下限， $Y^+$  与  $Y^-$  同理（以图 4-14 中的直方图为例，有  $X^+ = 1.2$ ,  $X^- = -0.6$ ,  $Y^+ = 0.6$ ,  $Y^- = -0.4$ ）。

最后，将  $\hat{X}_{orth}^U$  与  $\hat{Y}_{orth}^U$  代入式 (4.19) 中，即可得到未知光源  $U$  在相机 RGB 颜色空间中的响应值估计  $\hat{D}_r^U$ 、..，进而得到图像  $\mathcal{I}$  对应的白平衡校正系数：

$$\begin{cases} \hat{\mathcal{G}}_r^{\mathcal{I}} = \frac{1}{\hat{D}_r^U} \\ \hat{\mathcal{G}}_b^{\mathcal{I}} = \frac{1}{\hat{D}_b^U} \end{cases} \quad (4.27)$$

- 导致中性色像素统计算法失效的特殊情况

候选中性色像素直方图  $\mathbf{H}_n$  中含有足够的像素数是保证中性色像素统计算法能够可靠工作的必要条件之一。然而，存在以下两类特殊情况有可能导致  $\mathbf{H}_n$  中的像素数过少甚至成为空集：

- 1) 拍摄场景的照明光源色品位于中性色区域之外；
- 2) 场景中无中性色表面。

显然，当情况 1 发生时，使用式 (4.25) 进行光源颜色估计永远不可能得到正确的结

果，因为候选中性色像素直方图  $\mathbf{H}_n$  的重心必然只能位于中性色区域之内。然而，得益于正交色度平面的特性，当此类情况发生时，可以很容易地通过对其进行上、下、左、右四个方向的平移而使得候选中性色像素直方图  $\mathbf{H}_n$  中含有足够数量的像素，且这种平移恰好意味着对图像进行色温方向或沿着等色温线的方向的颜色调节。例如，在预先设定的中性色区域仅覆盖最低色温为 3000K 的光源的情况下，对于一张实际光源色温为 2800K 的图像，可通过对其亮度加权直方图  $H$  进行左移而使得估计出的光源色温在 3000K 附近。此时，用一组 3000K 光源对应的白平衡增益系数对 2800K 光源下拍摄的图像进行白平衡校正，仍然能够确保最终的输出结果处于一个尚可接受的范围之内。

当情况 2 发生时（对于日常摄影应用，这一事件发生的概率并不大，文献<sup>[163]</sup>中证实了至少有 80% 的日常场景中含有中性色物体表面），中性色像素统计算法将完全失效，此时只能求助于某些基于学习的模式识别算法，或者利用其它硬件（如额外的光度色度传感器等）进行辅助判断。然而，当候选中性色像素直方图  $\mathbf{H}_n$  中的像素数不足或是空集时，并无法判断该图像究竟是对应于情况 1 或是情况 2，此时只能按照情况 1 中对亮度加权直方图进行平移的做法对其进行处理。所幸的是，利用下文中即将介绍的色域映射算法，可以在一定程度上缓解由于情况 2 出现而导致的光源颜色估计错误。

#### 4.3.1.3 色域映射算法与中性色区域筛选

与中性色像素统计算法不同，色域映射 (gamut mapping) 算法无需对图像中的颜色分布提出任何先验假设，而是通过对输入图像的色域范围进行分析从而得到该场景可能对应的所有光源的集合。传统的色域映射算法<sup>[52, 53]</sup>通常采用某种预设策略从所有可能的光源的集合中选取一个作为其光源估计的结果，例如，使用该集合的几何中心作为光源的色度坐标。然而，在无额外约束的情况下，我们并无法证明该光源与真实光源相吻合的概率大于集合中的其它任一光源，换句话说，这种选取策略与在集合中进行随机选取并无本质区别。针对这一问题，本论文对传统的色域映射算法进行了改进，在得到可能的光源的集合后，并不尝试从中选取出唯一的光源，而是将该集合映射至正交色度平面中并与用户预先指定的中性色区域  $\mathbf{M}_n$  计算交集，从而筛选出一个面积更小的中性色区域，进一步提高中性色像素统计算法的准确性。

由式 (3.66) 可知，物体表面的反射特性、照明光源的光谱特性以及数码相机的响应特性共同决定了相机对于该表面所记录的原始响应值。给定一幅输入图像  $\mathcal{I}$ ，我们将颜色空间中能够包围其所有像素响应值构成的散点集合的最小凸集称为该图像的色域。以 RGB

图像为例，该图像中每个像素的响应值对应了 RGB 颜色空间中的一个坐标点，因此对所有像素对应的散点求得的最小凸多面体即为该图像在 RGB 颜色空间中的色域，如图 4-15 所示。

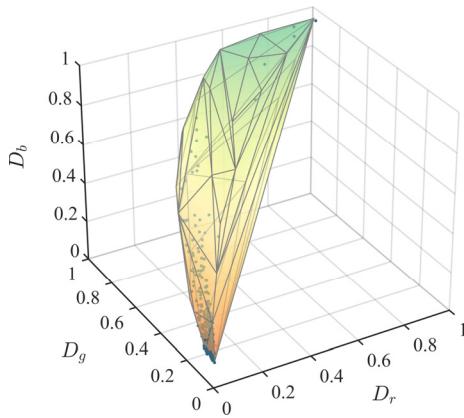


图 4-15 Nikon D3x 拍摄的测试图像（见图 4-13）在相机 RGB 颜色空间中对应的色域多面体

对于大部分日常场景下拍摄的图像来说，其色域范围总是十分有限的，例如图 4-15 中，色域多面体仅仅占据整个 RGB 颜色空间中的一小部分。基于类似的观察，色域映射算法认为，任何图像所包含的颜色总是有限的，其边界除了取决于场景中物体的颜色特性之外，还取决于该场景照明光源的颜色特性，因此图像的色域边界实际上为照明光源的颜色推断提供了一定的线索。例如，在一个高色温的光源下，无论是人眼还是相机，都无法观察到高饱和的红色或橙色。由此可知，倘若在一幅图像中观察到了高饱和的红色像素，那么该场景的照明光源一定不属于高色温光源；同理，若在一幅图像中观察到了高饱和的蓝色像素，则我们能够将低色温光源从其备选光源中排除出去。

- 算法原理

首先考虑 RGB 三维颜色空间中的情况。

我们将未知光源  $U$  下拍摄的输入图像  $\mathcal{I}$  的色域（gamut of image）记为  $T^U(\mathcal{I})$ ，将测试相机的标准色域（standard gamut）<sup>\*</sup>记为  $T^C$ ，并分别使用  $N^U$  与  $N^C$  表示其（多面体）顶点个数。假设利用一组白平衡增益系数  $\mathcal{G}_r$ 、 $\mathcal{G}_g$ 、 $\mathcal{G}_b$  能够将未知光源  $U$  下拍摄的图像校正至参考光源  $C$  下的颜色外貌，则必然有

---

<sup>\*</sup> 相机的标准色域指的是在颜色空间（或色度平面）中能够包围该相机对于所有可能存在的物体表面在参考光源下的响应值坐标的最小凸多面体（或多边形）。本论文基于响应值预测模型对标准色域  $T^C$  进行计算，其具体原理及实现过程请参阅第 5.1.3 小节。

$$\mathbf{G} \cdot \mathbf{D} \in \mathcal{T}^C, \quad \forall \mathbf{D} \in \mathcal{T}^U(\mathcal{I}) \quad (4.28)$$

其中

$$\mathbf{G} = \begin{bmatrix} \mathcal{G}_r & & \\ & \mathcal{G}_g & \\ & & \mathcal{G}_b \end{bmatrix} \quad (4.29)$$

表示由白平衡增益系数构成的 $3 \times 3$ 对角矩阵， $\mathbf{D} = [D_r, D_g, D_b]^T$ 表示图像 $\mathcal{I}$ 中某个像素的响应值。

由于 $\mathcal{T}^U(\mathcal{I})$ 与 $\mathcal{T}^C$ 均是外凸的，因此在研究映射问题时仅需考虑位于多面体边界上的那些响应值。对于多面体 $\mathcal{T}^U(\mathcal{I})$ 上的第*i*个顶点 $\mathbf{V}_i^U = [D_{i,r}^U, D_{i,g}^U, D_{i,b}^U]^T$ ，有且仅有 $N^C$ 个 $3 \times 3$ 对角矩阵 $\mathbf{G}$ 能够将其恰好映射至 $\mathcal{T}^C$ 的各个顶点，其中第*j*个对角矩阵（将 $\mathbf{V}_i^U$ 映射至 $\mathcal{T}^C$ 上第*j*个顶点的矩阵） $\mathbf{G}_{ij}$ 可表示为

$$\mathbf{G}_{ij} = \begin{bmatrix} D_{j,r}^C / D_{i,r}^U & & \\ & D_{j,g}^C / D_{i,g}^U & \\ & & D_{j,b}^C / D_{i,b}^U \end{bmatrix}, \quad 1 \leq i \leq N^U, 1 \leq j \leq N^C \quad (4.30)$$

显然，在由这 $N^C$ 个映射构成的集合中，任意数量元素之间的非负线性组合所产生的新的对角矩阵，也都能够将 $\mathbf{V}_i^U$ 映射至 $\mathcal{T}^C$ 之内\*，因此本论文将 $\mathcal{G}_r \mathcal{G}_g \mathcal{G}_b$ 空间中包围这 $N^C$ 个映射的最小凸多面体称为 $\mathbf{V}_i^U$ 至 $\mathcal{T}^C$ 的顶点映射域（gamut of vertex mappings），并使用符号 $\mathcal{T}(\mathcal{T}^C / \mathbf{V}_i^U)$ 进行表示：

$$\mathbf{G} \cdot \mathbf{V}_i^U \in \mathcal{T}^C, \quad \forall \mathbf{G} \in \mathcal{T}(\mathcal{T}^C / \mathbf{V}_i^U) \quad (4.31)$$

同理，对于多面体 $\mathcal{T}^U$ 的其它顶点，也各自对应着一个顶点映射域。对这 $N^U$ 个映射域计算交集，就得到了一个能够将多面体 $\mathcal{T}^U(\mathcal{I})$ 的每个顶点都映射至多面体 $\mathcal{T}^C$ 内部的映射域：

$$\mathcal{T}[\mathcal{T}^C / \mathcal{T}^U(\mathcal{I})] = \bigcap_{i=1}^{N^U} \mathcal{T}(\mathcal{T}^C / \mathbf{V}_i^U) \quad (4.32)$$

自然地，该映射域中的每一组映射，也能够将多面体 $\mathcal{T}^U(\mathcal{I})$ 自身映射至多面体 $\mathcal{T}^C$ 内

\* 严格来说，应该是“不位于多面体之外”。

部，即：

$$\mathbf{G} \cdot \mathbf{D} \in \mathcal{T}^C, \quad \forall \mathbf{G} \in \mathcal{T} \left[ \mathcal{T}^C / \mathcal{T}^U(\mathcal{I}) \right] \text{ and } \forall \mathbf{D} \in \mathcal{T}^U(\mathcal{I}) \quad (4.33)$$

由于映射域  $\mathcal{T} \left[ \mathcal{T}^C / \mathcal{T}^U(\mathcal{I}) \right]$  表征的是两个色域之间的映射关系，我们称之为色域映射域（gamut of gamuts mappings）。

显然，位于色域映射域  $\mathcal{T} \left[ \mathcal{T}^C / \mathcal{T}^U(\mathcal{I}) \right]$  内的每一组映射，都有可能将图像  $\mathcal{I}$  校正至参考光源  $C$  下的颜色外貌；反之，位于  $\mathcal{T} \left[ \mathcal{T}^C / \mathcal{T}^U(\mathcal{I}) \right]$  之外的映射，一定不可能对图像  $\mathcal{I}$  进行正确的白平衡校正。

需要注意的是，本小节中，我们使用“映射的集合”表示将  $\mathbf{V}_i^U$  恰好映射至  $\mathcal{T}^C$  顶点的  $N^C$  个映射，而使用“顶点映射域”表示所有能够将  $\mathbf{V}_i^U$  映射至  $\mathcal{T}^C$  之内的无穷多组映射的总体；同理，当我们使用“色域映射域”的说法时，指代的是满足式 (4.33) 的无穷多组映射构成的整体。此外，理解色域映射算法的核心在于对“图像的色域（gamut of image）”与“映射域（gamut of mappings）”加以区分，前者中的每个元素代表了一组相机响应值，而后者中每个元素代表的是一组从  $U$  光源至  $C$  光源的映射关系。

- 二维色域映射算法

在三维 RGB 颜色空间中实施色域映射算法的一个主要问题在于，一旦图像  $\mathcal{I}$  中存在饱和像素，则使用式 (4.32) 进行映射域求交计算时极有可能得到空集。考虑到这一问题，本论文中，我们在  $\left( \frac{D_r}{D_g}, \frac{D_b}{D_g} \right)$  平面（下文简称为  $rb$  平面）内对图像色域  $\mathcal{T}^U(\mathcal{I})$  以及标准色域  $\mathcal{T}^C$  进行计算，此时，两者均为凸多边形，如图 4-16 所示。

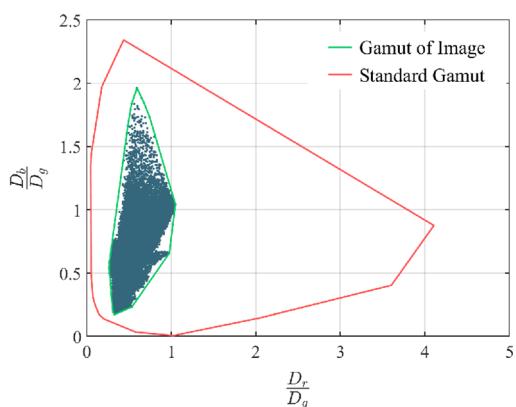


图 4-16  $rb$  平面上的标准色域  $\mathcal{T}^C$ （红色实线）与测试图像（见图 4-13）的图像色域  $\mathcal{T}^U(\mathcal{I})$ （绿色实线）示意图

之所以选择在  $rb$  平面上而非正交色度平面中对图像色域以及标准色域进行计算的原因有二：第一，由于变换公式 (4.18) 中涉及对数运算，三维 RGB 空间中色域的外凸性 (convexity) 有可能在变换至正交色度平面的过程中被破坏，无法满足色域映射算法的先决条件；第二，在  $rb$  平面上，一组  $\left(\frac{D_r}{D_g}, \frac{D_b}{D_g}\right)$  坐标至另一组  $\left(\frac{D_r}{D_g}, \frac{D_b}{D_g}\right)$  坐标的映射对应于一组响应值“调整”系数，而在正交色度平面中，一组  $(X_{orth}, Y_{orth})$  坐标至另一组  $(X_{orth}, Y_{orth})$  坐标的映射并不具有明确的物理含义。

在  $rb$  平面上，多边形  $T^U(\mathcal{I})$  上的第  $i$  个顶点的坐标可表示为  $\mathbf{V}_i^U = \left[ \frac{D_{i,r}^U}{D_{i,g}^U}, \frac{D_{i,b}^U}{D_{i,g}^U} \right]^\top$ ，此时，

式 (4.30) 需改写为

$$\mathbf{G}_{ij} = \begin{bmatrix} \frac{D_{j,r}^C D_{i,g}^U}{D_{j,g}^C D_{i,r}^U} \\ \frac{D_{j,g}^C D_{i,r}^U}{D_{j,r}^C D_{i,g}^U} \\ \frac{D_{j,b}^C D_{i,g}^U}{D_{j,g}^C D_{i,b}^U} \\ \frac{D_{j,g}^C D_{i,b}^U}{D_{j,b}^C D_{i,g}^U} \end{bmatrix}, \quad 1 \leq i \leq N^U, 1 \leq j \leq N^C \quad (4.34)$$

不妨将映射域  $T[\mathcal{T}^C / T^U(\mathcal{I})]$  中的元素坐标用  $(M_r, M_b)$  进行表示。根据式 (4.32)， $T[\mathcal{T}^C / T^U(\mathcal{I})]$  可通过对  $N^U$  个多边形（每个多边形对应着一个顶点映射域）求交得到，如图 4-17 所示。

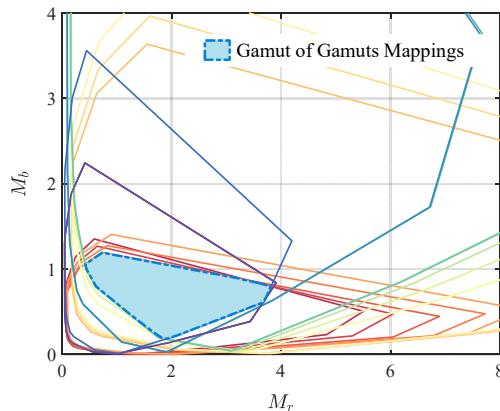


图 4-17 Nikon D3x 测试图像（参见图 4-13）的各个顶点映射域  $T(\mathcal{T}^C / \mathbf{V}_i^U)$ （彩色多边形）以及求交后的色域映射域  $T[\mathcal{T}^C / T^U(\mathcal{I})]$ （蓝色阴影）

既然任意一组  $(M_r, M_b) \in T[\mathcal{T}^C / T^U(\mathcal{I})]$  都能够将  $T^U(\mathcal{I})$  映射至  $\mathcal{T}^C$  内部，那么自然

地，也有可能将 $\left(\frac{D_{n,r}^U}{D_{n,g}^U}, \frac{D_{n,b}^U}{D_{n,g}^U}\right)$ 映射至 $\left(\frac{D_{n,r}^C}{D_{n,g}^C}, \frac{D_{n,b}^C}{D_{n,g}^C}\right)$ ，因此在 $\mathcal{T}\left[\mathcal{T}^C/\mathcal{T}^U(\mathcal{I})\right]$ 已知的情况下，通过计算逆映射，便可获得所有可能的 $\left(\frac{D_{n,r}^U}{D_{n,g}^U}, \frac{D_{n,b}^U}{D_{n,g}^U}\right)$ 坐标在 $rb$ 平面构成的光源色域（gamut of candidate illuminants） $\mathcal{T}(U)$ ：

$$\mathcal{T}(U) = \left\{ \left( \frac{D_{n,r}^C}{M_r D_{n,g}^C}, \frac{D_{n,b}^C}{M_b D_{n,g}^C} \right) \mid (M_r, M_b) \in \mathcal{T}\left[\mathcal{T}^C/\mathcal{T}^U(\mathcal{I})\right] \right\} \quad (4.35)$$

最后，将光源色域从 $rb$ 平面变换至正交色度平面中并与用户预先设定的候选中性色区域 $\mathcal{T}_n$ 求交，即可得到经色域映射算法筛选后的中性色区域 $\mathcal{T}_n^\cap$ ：

$$\mathcal{T}_n^\cap = \mathcal{T}_n \cap \mathcal{T}(U) \quad (4.36)$$

为了与候选中性色区域 $\mathcal{T}_n$ 加以区分，本论文将筛选后的中性色区域 $\mathcal{T}_n^\cap$ 称为“交集中性色区域（intersected neutral region）”。图 4-18 展示了测试图像在正交色度平面中对应的光源色域与交集中性色区域。

在得到交集中性色区域 $\mathcal{T}_n^\cap$ 后，可利用其生成对应的交集中性色区域掩模 $\mathbf{M}_n^\cap$ ，并以此替代式 (4.23) 中的候选中性色区域掩模 $\mathbf{M}_n$ 对交集中性色像素直方图（intersected neutral histogram）进行计算：

$$\mathbf{H}_n^\cap = \mathbf{H} \circ \mathbf{M}_n^\cap \quad (4.37)$$

后续过程与式 (4.25) ~ (4.27) 类似，不再赘述。

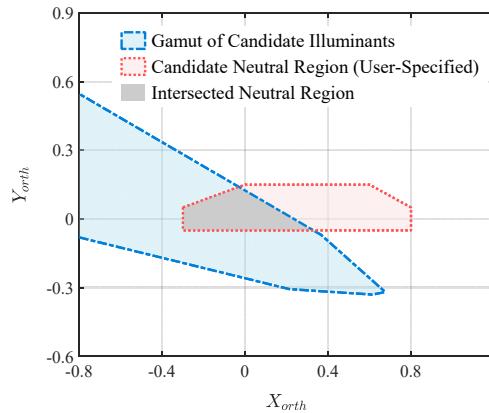


图 4-18 Nikon D3x 测试图像（参见图 4-13）在正交色度平面中的光源色域 $\mathcal{T}(U)$ （蓝色阴影）、用户指定的候选中性色区域 $\mathcal{T}_n$ （红色阴影），以及两者求交后的交集中性色区域 $\mathcal{T}_n^\cap$ （灰色阴影）

需要注意的是，虽然图像色域  $T^U(\mathcal{I})$  与光源色域  $T(U)$  中的各元素均表示  $rb$  平面上一组  $\left(\frac{D_r}{D_g}, \frac{D_b}{D_g}\right)$  坐标，然而前者表征的是由输入图像中  $\mathcal{I}$  所有像素所构成的色域范围，而后者仅仅表示光源  $U$  可能的色域范围。

#### 4.3.1.4 实验设置

本论文使用 Nikon D3x 数码相机在室内、外不同场景下使用不同的参数拍摄了共 134 张 raw 图像，并以此作为测试数据集对 GMap-NPStat 算法的性能以及使用原始响应值预测模型计算标定参数的可行性进行评估。为了获取拍摄场景中真实的光源颜色信息以便对光源颜色估计算法的准确性进行评估，我们在每个拍摄场景中均放置了一张 X-Rite Classic 标准色卡，如图 4-19 所示\*。

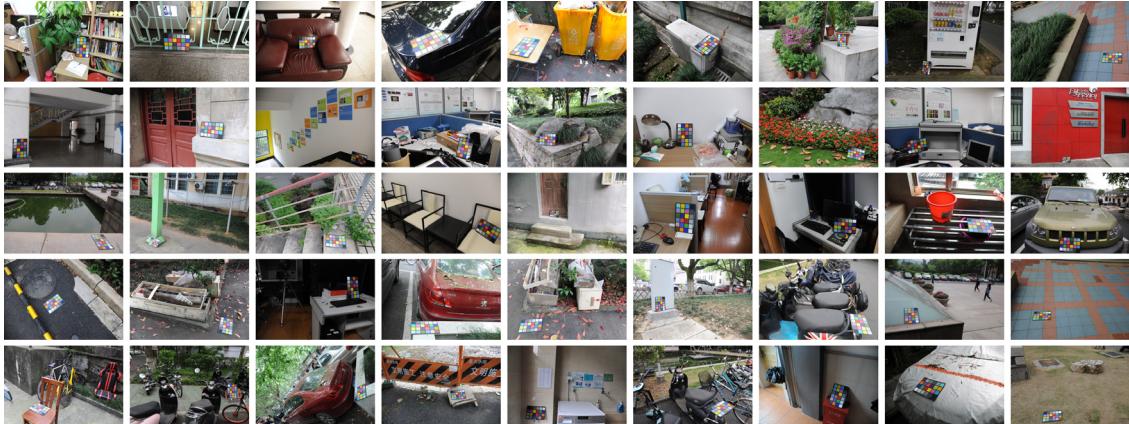


图 4-19 Nikon D3x ColorChecker 图像数据集中的部分图像

评估自动白平衡算法的性能时通常使用光源真实的 RGB 响应值与算法估计响应值之间的角度差作为光源颜色估计准确度的衡量指标<sup>[158]</sup>，即

$$\theta = \cos^{-1} \frac{\mathbf{D}^U \cdot \hat{\mathbf{D}}^U}{\|\mathbf{D}^U\| \|\hat{\mathbf{D}}^U\|} \quad (4.38)$$

式中  $\mathbf{D}^U = [D_r^U, D_g^U, D_b^U]$  与  $\hat{\mathbf{D}}^U = [\hat{D}_r^U, \hat{D}_g^U, \hat{D}_b^U]$  分别表示光源  $U$  的真实响应值与估计响应值。由于角度差指标仅取决于  $\mathbf{D}^U$  与  $\hat{\mathbf{D}}^U$  向量在空间中的方向而非模长，因此可以直接使用场景中的中性色物体表面的响应值对角度差进行计算。在数据处理过程中，我们手动地对标准色卡中的中性色块的响应值进行提取，并利用式 (4.38) 对每一张测试图像的光源颜

\* 为了更好地展示图像内容，图 4-19 中列出的均为 ISP 输出图像，并非是自动白平衡模块的输入 raw 图像。

色估计准确度进行评估。此外，在对测试图像执行光源颜色估计算法之前，我们预先使用一个黑色的掩模将标准色卡所在的区域进行遮盖，如图 4-20 所示，以确保标准色卡中的像素信息未参与任何决策过程。最后，为了避免输入图像中存在的部分彩色噪声对图像色域  $T^U(\mathcal{I})$  的计算产生干扰，我们预先使用  $\sigma=1.5$  的  $5 \times 5$  高斯卷积核对输入图像进行平滑处理，以提高色域映射算法的鲁棒性。



图 4-20 左：包含标准色卡测试 raw 图像；右：自动白平衡校正模块的输入图像，其中标准色卡所在区域已使用黑色掩模进行遮盖

#### 4.3.1.5 结果与分析

在对自动白平衡算法的性能进行评估时，除了使用角度差的均值作为评价指标外，通常还需要考察算法对于某些极端场景（hard images）的容错能力<sup>[164]</sup>。为此，本论文使用角度差的均值、中位数、三均值（trimean）、最好 25% 均值（best 25%）以及最差 25% 均值（worst 25%）共五种度量方式对 GMap-NPStat 算法在 Nikon D3x ColorChecker 数据集上的光源颜色估计准确度进行评价，其中，三均值、最好 25% 均值以及最差 25% 均值的计算公式分别为

$$\text{trimean}(\theta_1, \dots, \theta_N) = \frac{Q_1 + 2Q_2 + Q_3}{4} \quad (4.39)$$

$$\text{best 25\%}(\theta_1, \dots, \theta_N) = \frac{4}{N} \sum_{\theta_i < Q_1} \theta_i \quad (4.40)$$

以及

$$\text{worst 25\%}(\theta_1, \dots, \theta_N) = \frac{4}{N} \sum_{\theta_i > Q_3} \theta_i \quad (4.41)$$

式中  $Q_1$ 、 $Q_2$ 、 $Q_3$  分别表示角度差集合  $\{\theta_1, \dots, \theta_N\}$  中的第一、第二、第三四分位数。

表 4-2 中列出了 GMap-NPStat 算法与一些常见的基于先验假设的自动白平衡算法在 Nikon D3x ColorChecker 图像数据集上的光源颜色估计准确性比较结果，其中在 Weng 等

人提出的算法中<sup>[48]</sup>, 用于白点检测的阈值等参数均已针对 Nikon D3x 的设备特性进行了调整。从表 4-2 中的结果可以看出, 相对于传统的统计类自动白平衡算法, GMap-NPStat 算法无论是在光源颜色估计的整体性能(均值、中位数)或是针对极端场景的适应能力(最差 25% 均值)上均具备一定的优势。考虑到若能够在进行光源颜色估计之前预先对场景类别加以区分从而缩小预设的中性色区域  $T_n$  的范围, GMap-NPStat 算法的性能仍然存在进一步提升的空间。

表 4-2 不同算法在 Nikon D3x ColorChecker 图像数据集上的光源颜色估计准确度比较

算法	均值	中位数	三均值	最好 25% 均值	最差 25% 均值
White Patch <sup>[165]</sup>	7.13	4.27	5.53	1.03	16.49
1st Order Gray edge <sup>[43]</sup>	4.86	3.27	3.70	0.67	11.81
2nd Order Gray edge <sup>[43]</sup>	4.79	3.23	3.44	0.81	11.50
Gray World <sup>[166]</sup>	4.67	4.25	4.32	1.31	8.84
Shade of Gray <sup>[167]</sup>	4.37	3.24	3.57	0.65	10.01
Weng 等人 <sup>[48]</sup>	3.42	2.65	2.75	0.57	7.68
GMap-NPStat	3.01	2.37	2.62	0.63	6.35

图 4-21 中对部分测试图像的白平衡校正结果进行了比较, 为了更好地进行可视化展示, 其中所有图像均已利用相同的颜色校正矩阵<sup>[168]</sup>与非线性变换函数<sup>[169]</sup>完成了相机 RGB 颜色空间至 sRGB 颜色空间的转换。需要指出的是, 得益于正交色度平面的颜色分布特性, 即使当图像中几乎不存在中性色像素或中性色像素数量极少时, 也能够通过在  $(X_{orth}, Y_{orth})$  平面内对亮度加权直方图进行平移从而获得准确度尚可的光源颜色估计结果(如图 4-21 中第四列的场景)。

除了考察算法在图像数据集上的总体表现之外, 我们同时也针对个别的 GMap-NPStat 算法估计结果不够理想的场景进行了分析。图 4-22 列出了在 Nikon D3x ColorChecker 图像数据集中部分导致 GMap-NPStat 算法出现误判的测试图像及其相关的直方图统计结果。对于图 4-22 (a) 中的测试图像, 由于场景中包含了大量的伪中性色(草地中的淡黄色像素), 因此二维直方图向  $X_{orth}$  轴的正方向出现偏移, 导致最终估计出的光源色温明显低于真实的光源色温; 对于图 4-22 (b) 中的测试图像, 由于场景中存在饱和度较高的绿色像素,

导致色域映射算法计算得到的光源色域  $\mathcal{T}(U)$  未能包含真实的光源色品坐标，因此得到了错误的交集中性色区域  $\mathcal{T}_n^{\cap}$ ，这一问题可通过对标准色域  $\mathcal{T}^C$  进行重新定义而得到改善；对于图 4-22(c) 中的测试图像，由于场景中存在两种色温差异显著的光源（图像左侧主要为室外的自然光源照明，右侧主要为走道内的荧光灯照明），因此亮度加权直方图中存在两个明显的尖峰，此时利用式 (4.25) 对光源颜色进行估计时将不可避免地出现偏差。



图 4-21 部分测试图像的白平衡校正结果比较，其中第一行为手动白平衡校正结果 (ground-truth)，第二至七行分别为 1st Order Gray Edge 算法、2nd Order Gray Edge 算法、Gray World 算法、Shade of Gray 算法、Weng 等人的算法以及 GMap-NPStat 算法校正结果以及各自对应的光源颜色估计角度差数据（标准色卡所在的图像区域均未参与计算）\*

\* 图 4-21 中展示的均为 ISP 输出图像。除了白平衡模块采用各算法估计出的光源颜色作为校正依据外，其余各模块均为手动校正。

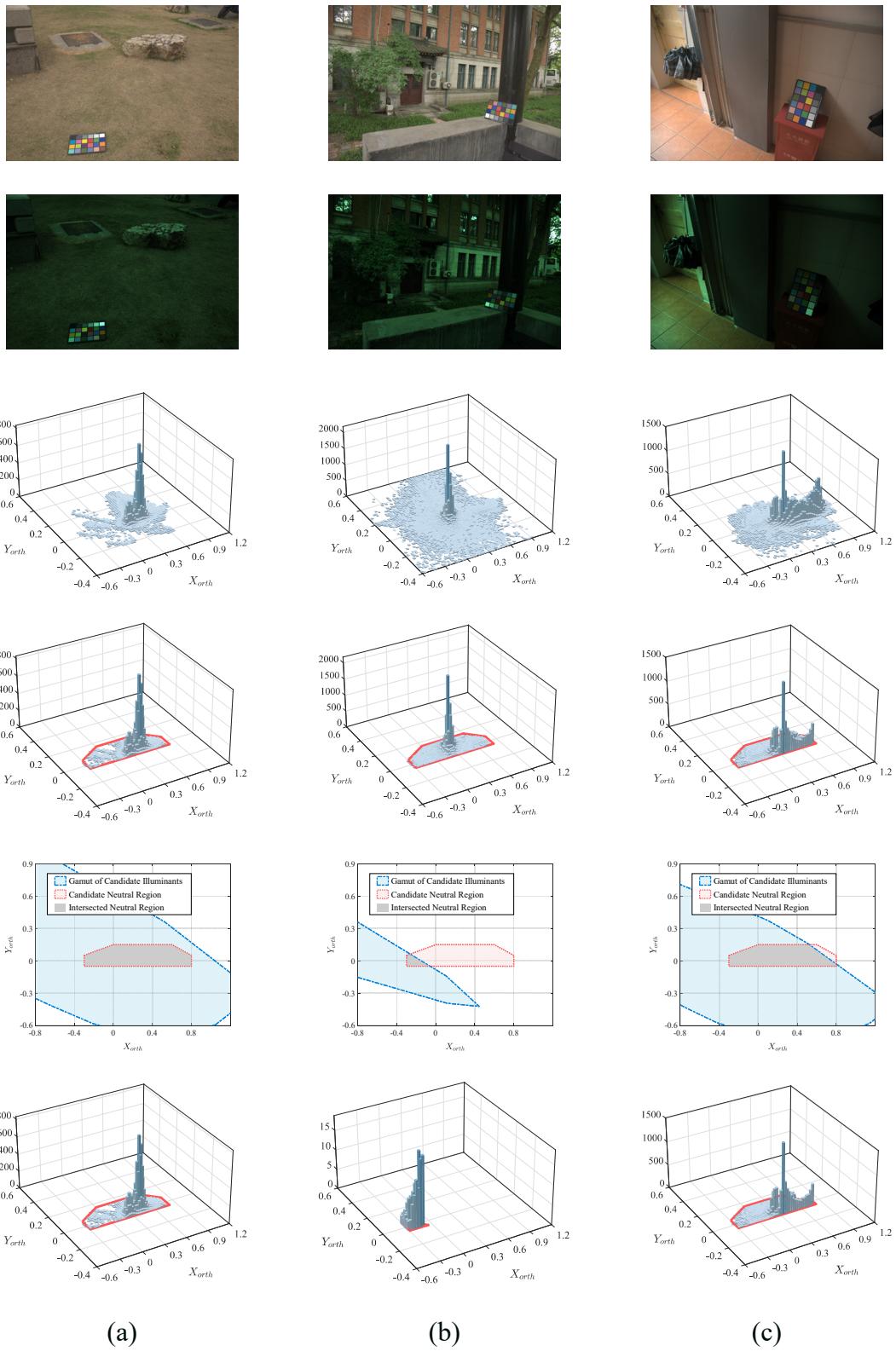


图 4-22 部分导致 GMap-NPStat 算法出现误判的测试图像及其相关参数。由上至下：手动白平衡校正结果（ISP 输出图像）、自动白平衡模块的输入 raw 图像、亮度加权直方图  $\mathbf{H}$ 、从候选中性色区域  $T_n$  中生成的候选中性色像素直方图  $\mathbf{H}_n$ 、通过色域映射算法得到的交集中性色区域  $T_n^\cap$  以及从交集中性色区域  $T_n^\cap$  中生成的交集中性色像素直方图  $\mathbf{H}_n^\cap$

### 4.3.2 基于特征图二次卷积的自动白平衡算法（ReWNet）

除了基于一些先验假设对图像进行分析之外，许多研究者也尝试着使用学习的方法对光源颜色回归模型中的参数进行计算<sup>[54, 56, 170-172]</sup>。近年来，随着深度学习技术在计算机视觉领域的大放异彩，一些研究者也尝试着将其中的算法思路与模型架构用于颜色恒常性计算领域，并在基准数据集上取得了前所未有的性能提升<sup>[173-176]</sup>。然而，一些用于光源颜色估计的深度模型仅仅是在已有的图像分类模型的基础上进行了少许改动<sup>[59, 177]</sup>，其工作原理以及特征提取机制尚未得到很好的解释，且参数规模往往过于庞大<sup>[177, 178]</sup>，并不适用于白平衡校正这一应用场景。

本小节中，针对现有的基于深度网络的自动白平衡算法存在的若干不足，我们提出了一种具有高可解释性的卷积神经网络模型——二次卷积网络（reweight network，下文中简称为 ReWNet）。该模型利用二次卷积单元（reweight unit）对传统自动白平衡算法中的“有效像素筛选”这一概念进行了抽象表征，允许模型在训练过程中同时对多种能够为光源颜色估计提供线索的颜色特征进行学习（例如肤色、植物的绿色等能够作为场景色温判断依据的典型颜色）。在达到相近的估计准确度的前提下，ReWNet 仅需要现有的深度模型 1% ~ 3% 的参数规模以及 8% ~ 20% 的计算量，因此十分适合作为轻量化模型部署于智能终端或车载成像设备等对于计算量与功耗具有较高限制的移动平台之中。

#### 4.3.2.1 模型与方法

我们首先从网络的结构单元与整体架构两个层面对 ReWNet 的设计思路以及工作原理进行介绍。

ReWNet 的主体部分由传统的卷积单元（convolution block）与特征图二次卷积单元（feature map reweight unit，下文中简称为 ReWU）共同组成，其中卷积单元的工作原理与普通的卷积神经网络无异，即通过逐层的卷积操作从而实现图像不同层级的特征提取；ReWU 则在卷积单元提取出的特征图的基础上再次进行权重分配，从而蒸馏出对于光源颜色估计有益的特征信息，抑制其余无关信息。

不同于图像识别、理解等高级视觉任务，在光源颜色估计这一任务中，即使是层级较低的特征图（甚至是输入图像）中仍然包含了与光源颜色信息有关的重要线索。基于这一考虑，我们并未采用图像分类模型中常见的串联型网络架构<sup>[14, 179]</sup>，而是将来自不同层级的 ReWU 的输出向量共同汇总至一个多层次感知机（multilayer perceptron）中，并经由若干组全连接层（fully-connected layer）回归出光源颜色估计  $\hat{D}_r^U, \hat{D}_g^U, \hat{D}_b^U$ 。

- 特征图二次卷积单元 (ReWU)

传统的中性色像素搜索算法通常利用预先设定的一些约束条件对目标像素是否具有中性色属性进行判断<sup>[48, 157]</sup>, 例如在 4.3.1.2 小节中介绍的 NPStat 算法中, 为二维直方图设定候选中性色区域即是在对输入像素施加一定的约束。不妨考虑最简单的一种情况: 假设需要判断某一像素  $x$  的色品坐标是否位于中性点附近, 可使用如下判据:

$$x \triangleq \begin{cases} \text{neutral} & \text{if } |(u_x - u_0) - (v_x - v_0)| \leq T \text{ and } |(u_x - u_0) + (v_x - v_0)| \leq T \\ \text{chromatic} & \text{otherwise} \end{cases} \quad (4.42)$$

其中  $(u_x, v_x)$  与  $(u_0, v_0)$  分别表示该像素与理想中性色的色度平面坐标,  $T$  表示一个用于调整(近) 中性色区域大小的阈值。

为了使形如式 (4.42) 的条件表达式能够更加方便地参与到深度学习框架的张量运算之中, 将其改写为矩阵形式:

$$x \triangleq \begin{cases} \text{neutral} & \text{if } \mathcal{W}(u_x, v_x) > 0 \\ \text{chromatic} & \text{otherwise} \end{cases} \quad (4.43)$$

$$\mathcal{W}(u_x, v_x) = \text{ReLU} \left\{ \min \left( \mathbf{A} [u_x, v_x]^T + \mathbf{b} \right) \right\} \quad (4.44)$$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} T - u_0 + v_0 \\ T + u_0 - v_0 \\ T - u_0 - v_0 \\ T + u_0 + v_0 \end{bmatrix} \quad (4.45)$$

式中  $\min(\cdot)$  表示返回输入向量中最小的元素值,  $\text{ReLU}(\cdot)$  表示线性整流单元<sup>[180]</sup>, 其将输入向量中的所有负值元素进行截零操作。图 4-23 展示了使用式 (4.43) ~ (4.45) 对一张  $2 \times 2$  “图像”中两个近中性色像素进行提取的过程。

从这种类似于选择性注意力(selective attention)机制中得到了启发, 本论文将式 (4.43) ~ (4.45) 中的像素筛选过程进行了抽象化描述, 并在此基础上提出了特征图二次卷积单元 ReWU。每个 ReWU 以传统卷积单元输出的特征图 (RGB 原图像也可视作一种特殊的特征图) 作为输入, 通过对约束条件以及阈值大小进行学习从而实现对特征图中不同像素进行权重再分配。理论上, 我们希望 ReWU 能够强调特征图中那些对于光源颜色估计能够提供重要线索的像素, 同时抑制甚至完全忽略那些相对不重要的像素。

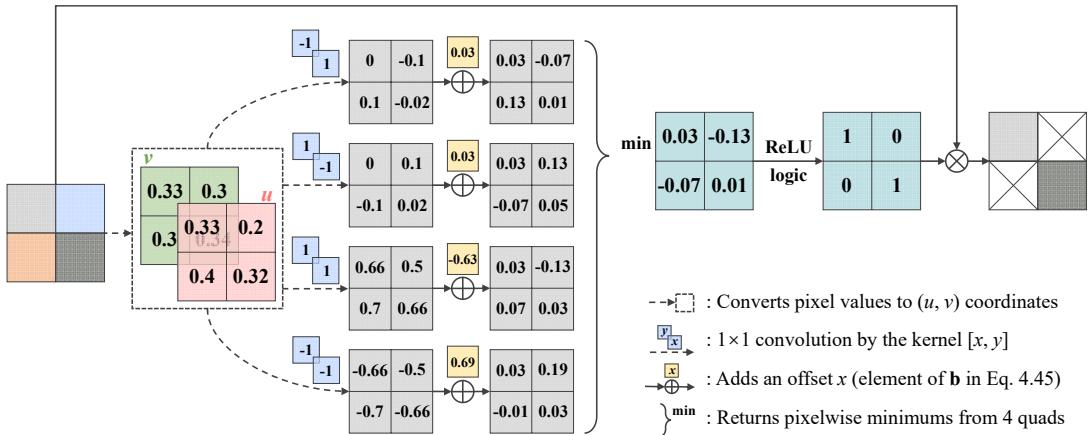


图 4-23 基于张量运算（参见式 (4.43) ~ (4.45)）的近中性色像素提取过程示意图。本例中有  $(u_0, v_0) = (0.33, 0.33)$  以及  $T = 0.03$

具体来说，给定一张输入特征图  $\mathbf{M} \in \mathbb{R}^{H \times W \times C}$ ，ReWU 将通过  $1 \times 1$  卷积、最小值搜索以及非线性激活共三步操作得到一张权重分配图（reweighting map） $\mathbf{W} \in \mathbb{R}^{H \times W}$ ，并将其与输入特征图的逐像素乘积（pixel-wise multiplication）作为输出特征图  $\mathbf{M}' \in \mathbb{R}^{H \times W \times C}$ ：

$$\mathbf{M}' = \mathbf{W} \circ \mathbf{M} \quad (4.46)$$

$$\mathbf{W} = \kappa \cdot \text{ReLU} \left[ \min(\mathbf{g} * \mathbf{M} + \mathbf{T}) \right] \quad (4.47)$$

式中  $\mathbf{g}$  表示一个  $1 \times 1 \times C \times K$  张量， $C$  代表输入特征图  $\mathbf{M}$  的通道数， $K$  代表卷积核的个数（某种意义上也可理解为对  $\mathbf{M}$  中像素所施加的约束的个数）， $\mathbf{T}$  表示一个用于存放阈值参数的  $1 \times K$  向量\*，符号“ $*$ ”表示卷积操作， $\min(\cdot)$  表示对张量沿通道方向计算最小值， $\kappa$  表示一个幅值调整系数。图 4-24 中展示了一个 ReWU 的基本结构组成。由于输入特征图  $\mathbf{M}$  在进入 ReWU 之前已经由传统卷积单元进行了一次“加权”操作，因此本论文将该权重重新分配的过程称为“二次卷积（rewriting）”。

显然，图 4-23 中的近中性色像素提取过程仅仅是 ReWU 的一个特殊范例，在该范例中，我们在二维色度平面上对输入特征图（即图像本身）共施加了 4 个约束条件，并使用布尔权重分配图（ $\mathbf{W}$  中的元素要么是 0 要么是 1）对输入特征图进行二次加权。相比于该

\* 本小节的公式中，当运算符号两侧的张量具有不同尺寸时，默认使用自动维度扩展机制（auto-broadcasting）对其进行尺寸匹配。例如，在式 (4.47) 中，加号左侧的张量 ( $\mathbf{g} * \mathbf{M}$ ) 的尺寸为  $H \times W \times K$ ，而右侧  $\mathbf{T}$  的尺寸为  $1 \times K$ ，此时，首先对  $\mathbf{T}$  进行维度重构，使其成为一个  $1 \times 1 \times K$  张量，并在前两个维度上对其进行复制填充，使其成为一个  $H \times W \times K$  张量，以便与张量 ( $\mathbf{g} * \mathbf{M}$ ) 进行逐像素相加运算（pixel-wise addition）。

范例，ReWU 对此类像素筛选机制进行了更加抽象的表征，其并未预先指定用于搜索目标像素的颜色空间，也并未对用于筛选目标像素的约束条件以及阈值做出任何预设，而是通过学习的方法在网络训练过程中不断地对其中的参数做出调整，从而获得与光源颜色估计有关的特征提取能力。

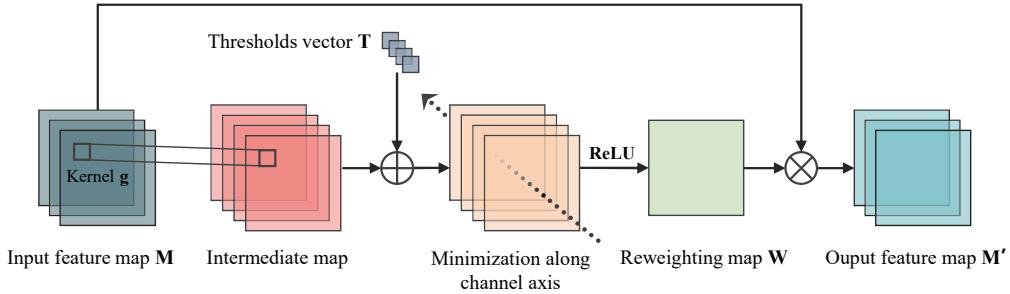


图 4-24 二次卷积单元 (ReWU) 的基本结构组成

本论文中，我们令每个 ReWU 中卷积核的个数  $K$  等于其输入特征图的通道数  $C$ ，因此单个 ReWU 仅仅需要  $O(C^2)$  的参数规模，较之于传统的卷积单元其空间复杂度降低了约 1~2 个数量级<sup>[14]</sup>。

- 二次卷积网络 (ReWNet) 的整体架构

ReWNet 的整体网络架构如图 4-25 所示。从数据的流动过程来看，ReWNet 的工作原理与传统的卷积神经网络类似，即通过序贯排列的卷积单元从输入图像中进行特征的逐层提取。然而，与传统的卷积神经网络不同的是，ReWNet 并未将卷积单元的输出直接传递至全连接层中，而是使用 ReWU 对每个卷积单元的输出特征图进行二次权重分配。此外，有别于传统卷积神经网络所采用的将最高层（即数据传输链的最后端） $H_N \times W_N \times C_N$  特征图“拍平 (flatten)”至  $H_N W_N C_N \times 1$  向量的方法，ReWNet 使用全局均值池化策略 (global average pooling)<sup>[181]</sup> 从每个 ReWU 的输出特征图  $\mathbf{M}'$  中生成一组长度仅为  $C_i$  的一维激活向量，并将其与来自其它 ReWU 的激活向量进行连结 (concatenation)，从而生成一组长度为  $3 + \sum_{i=1}^N C_i$  的向量并输入至后续的全连接层中\*。全局均值池化策略的使用在大幅减少全连

\*  $H_i$ 、 $W_i$ 、 $C_i$  分别表示传统的卷积神经网络中第  $i$  层卷积单元输出特征图的高、宽以及通道数。当一个卷积神经网络共包含  $N$  层卷积单元时， $H_N$ 、 $W_N$ 、 $C_N$  即表示最高层特征图的尺寸。对于 ReWNet，每个 ReWU 的输出特征图  $\mathbf{M}'$  与其输入特征图  $\mathbf{M}$ （亦为其前端卷积单元的输出特征图）具有相同的尺寸，因此第  $i$  层 ReWU 的输出特征图尺寸也为  $H_i \times W_i \times C_i$ 。

接层中神经元个数的同时，也使得一维激活向量的长度不受特征图尺寸的影响，这一性质确保了 ReWNet 能够接受任何尺寸的图像作为输入而无需对网络结构做出调整。

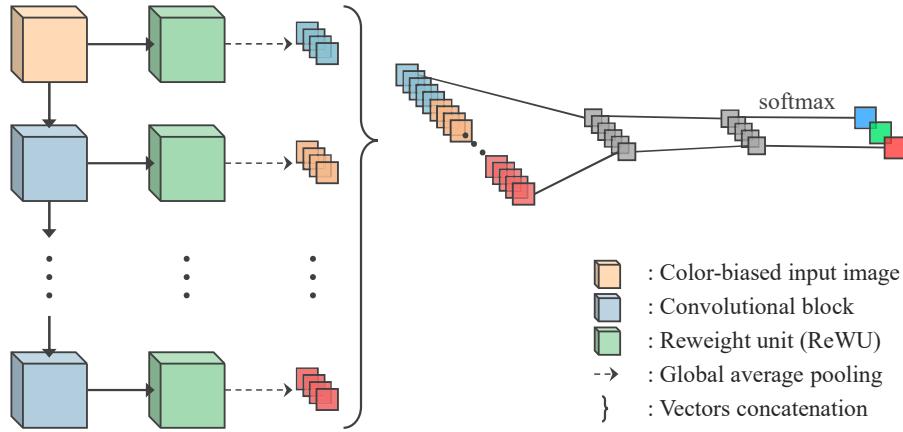


图 4-25 ReWNet 的整体网络架构

本论文分别对层级数  $N$  (即卷积单元的个数，或 ReWU 的个数减去 1) 为 1、2、3 的三种 ReWNet 实例 (下文分别称为 1-Hierarchy、2-Hierarchy 以及 3-Hierarchy 模型) 进行了测试与比较。对于 ReWNet 中的传统卷积单元，本论文选择使用 Inception v3<sup>[182]</sup>模型中的卷积核设计，并将前三层卷积单元中的卷积核的个数分别设为 32、32、64。对于除第 0 层级 (直接连接输入 RGB 图像的 ReWU) 之外的所有 ReWU，我们令其  $1 \times 1$  卷积核的个数与输入特征图的通道数相等，即  $K_i = C_i, i \in \{1, \dots, N\}$ ；而对于第 0 层级的 ReWU，我们令  $K_0 = 16$ 。

表 4-3 列出了三种网络实例中具体的层级结构设计以及参数选择。

表 4-3 具有不同层级数的 ReWNet 实例设计与参数选择

ReWNet 实例		1-Hierarchy		2-Hierarchy		3-Hierarchy	
网络层级 名称	输出特征图 尺寸	卷积单元或 全连接层中 核的尺寸	ReWU 中卷 积核的尺寸	卷积单元或 全连接层中 核的尺寸	ReWU 中卷 积核的尺寸	卷积单元或 全连接层中 核的尺寸	ReWU 中卷 积核的尺寸
输入层 (卷积层 0)	224×224	-	1×1, 16	224×224	-	1×1, 16	224×224
卷积层 1	112×112	3×3, 32, stride 2	1×1, 32	112×112	3×3, 32, stride 2	1×1, 32	112×112
卷积层 2	-	-	-	112×112	3×3, 32	1×1, 32	112×112
卷积层 3				-	-	-	112×112
连接层	35×1	-		67×1	-	-	131×1
全连接层 1	64×1	1, 64		128×1	1, 128		256×1
全连接层 2	32×1	1, 32		64×1	1, 64		128×1
全连接层 3	16×1	1, 16		32×1	1, 32		64×1
输出层	3×1	1, 3		3×1	1, 3		3×1
							1, 3

### 4.3.2.2 实验设置

- ReWNet 参数初始化

对于 ReWNet 中的传统卷积单元，我们直接使用针对 ImageNet 图像分类任务训练得到的 Inception v3 模型中的权重参数对其进行初始化；对于位于网络最后端多层感知机中的全连接层，我们使用满足正态分布的伪随机数对其进行初始化。

关于 ReWU 中卷积核张量  $\mathbf{g}$  以及阈值向量  $\mathbf{T}$  的初始化步骤相比传统卷积单元而言将略微复杂一些。不失一般性地，我们假设某一个 ReWU 的输入特征图  $\mathbf{M}$  中各像素之间独立同分布 (independent and identically distributed)，且像素值均由一个服从  $[0,1]$  区间内均匀分布的随机数生成器产生。此时，若使用来自标准正态分布的随机数对张量  $\mathbf{g}$  进行初始化并将向量  $\mathbf{T}$  中所有元素均初始化为 0，那么有

$$E(Z) = E(G)E(M) = 0 \quad (4.48)$$

$$\begin{aligned} \text{var}(Z) &= C \left[ \text{var}(G) \text{var}(M) + \text{var}(G)E^2(M) + \text{var}(M)E^2(G) \right] \\ &= C \left[ 1 \times 1/12 + 1 \times (1/2)^2 + 1/12 \times 0 \right] = \frac{C}{3} \end{aligned} \quad (4.49)$$

式中  $G$ 、 $M$ 、 $Z$  分别表示  $\mathbf{g}$ 、 $\mathbf{M}$  以及  $(\mathbf{g} * \mathbf{M})$  中的随机变量， $C$  表示特征图  $\mathbf{M}$  中的通道数， $E(\cdot)$  与  $\text{var}(\cdot)$  分别表示随机变量的数学期望与方差。显然， $(\mathbf{g} * \mathbf{M})$  中的随机变量  $Z$  将满足  $\mu = 0$ 、 $\sigma = \sqrt{C/3}$  的正态分布，因此使用  $\min(\cdot)$  沿通道方向计算其最小值必然得到小于 0 的值，导致经线性整流单元后得到的权重分配图  $\mathbf{W}$  中所有元素值全为 0。此时，ReWU 中的参数在训练过程中将无法得到有效的更新，模型势必无法收敛。

为解决这一问题，本论文提出了一种启发性的阈值向量  $\mathbf{T}$  的初始化方法，以确保对于任意的输入特征图  $\mathbf{M}$ ，初始条件下的权重分配图  $\mathbf{W}$  中恰好有一半的元素非零。

首先，仍然使用来自标准正态分布的随机数对张量  $\mathbf{g}$  进行初始化。在使用  $\mathbf{g}$  对输入特征图  $\mathbf{M}$  进行卷积操作之后，我们对  $(\mathbf{g} * \mathbf{M})$  中的各个像素独立地沿通道方向进行归一化，使得其中的随机变量具有 0 数学期望与单位标准差：

$$\hat{z}_{x,k} = \frac{z_{x,k} - \mu_x}{\sigma_x + \epsilon} \quad (4.50)$$

式中  $k \in \{1, \dots, K\}$  表示通道索引， $K$  表示.. 中的通道数， $z_{x,k}$  与  $\hat{z}_{x,k}$  分别表示归一化前后  $(\mathbf{g} * \mathbf{M})$  中的像素  $x$  在第  $k$  个通道上的响应值， $\epsilon$  为一个很小的常数用于确保分母不为零， $\mu_x$  与  $\sigma_x$  分别表示像素  $x$  在通道方向上的均值与标准差：

$$\mu_x = \frac{1}{K} \sum_{k=1}^K z_{x,k}, \quad \sigma_x = \sqrt{\frac{1}{K} \sum_{k=1}^K (z_{x,k} - \mu_x)^2} \quad (4.51)$$

我们将该归一化过程称为通道归一化 (channel normalization)，并用符号  $CN(\cdot)$  进行表示。图 4-26 中将通道归一化技术与其余三类深度学习中常用的特征图归一化技术进行了可视化比较。

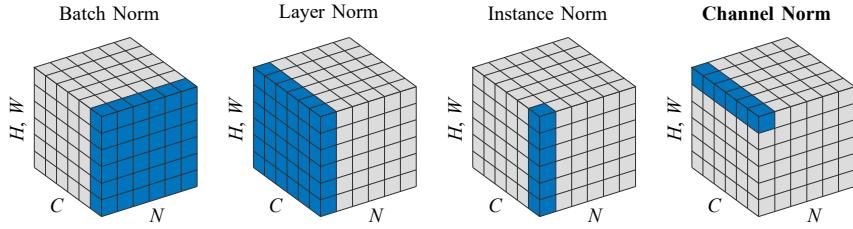


图 4-26 通道归一化 (Channel Norm) 与批归一化 (Batch Norm)<sup>[183]</sup>、层归一化 (Layer Norm)<sup>[184]</sup> 以及实例归一化 (Instance Norm)<sup>[185]</sup> 的比较。图中每个大立方体表示一个由若干特征图构成的张量， $N$  代表数据批 (batch) 维度， $C$  代表通道维度， $H$  与  $W$  代表特征图的空间维度。在进行归一化时，所有蓝色小立方体所对应的元素将共享一组相同的均值与标准差

加入通道归一化操作后，式 (4.47) 需改写为

$$\mathbf{W} = k \cdot \text{ReLU} \left\{ \min \left[ CN(\mathbf{g} * \mathbf{M}) + \mathbf{T} \right] \right\} \quad (4.52)$$

由随机变量的统计特性可知，在  $CN(\mathbf{g} * \mathbf{M})$  中各像素相互独立的情况下， $\min [CN(\mathbf{g} * \mathbf{M})]$  中的随机变量将近似满足对称分布，且具有数学期望：

$$E \left\{ \min \left[ CN(\mathbf{g} * \mathbf{M}) \right] \right\} = K \int_{-\infty}^{\infty} z \varphi(z; 0, 1) [1 - \Phi(z; 0, 1)]^{(K-1)} dz \quad (4.53)$$

式中  $\varphi(z; 0, 1)$  与  $\Phi(z; 0, 1)$  分别表示标准正态分布的概率分布函数 (PDF) 与累积分布函数 (CDF)。式 (4.53) 的详细推导过程请参见附录二。

为了使  $\min [CN(\mathbf{g} * \mathbf{M}) + \mathbf{T}]$  中的随机变量仍然能够近似满足对称分布且具有 0 数学期望，可使用常数  $\tau$  对阈值向量  $\mathbf{T}$  中的所有元素进行初始化，即

$$\mathbf{T} = \underbrace{[\tau, \tau, \dots, \tau]}_K \quad (4.54)$$

显然，为了使  $E \left\{ \min \left[ CN(\mathbf{g} * \mathbf{M}) + \mathbf{T} \right] \right\} = 0$  成立，应有

$$\begin{aligned}
\tau &= E(\tau) \\
&= -E \left\{ \min \left[ CN(\mathbf{g} * \mathbf{M}) \right] \right\} \\
&= -K \int_{-\infty}^{\infty} z \varphi(z; 0, 1) [1 - \Phi(z; 0, 1)]^{(K-1)} dz
\end{aligned} \tag{4.55}$$

此时，由于  $\min \left[ CN(\mathbf{g} * \mathbf{M}) + \mathbf{T} \right]$  中的随机变量具有 0 数学期望且近似满足对称分布，因此经线性整流单元 ReLU 截零后得到的权重分配图  $\mathbf{W}$  中，非零元素的个数将恰好等于  $\mathbf{W}$  总像素数的一半。

- 基准数据集与图像预处理

为了充分考察模型性能的设备相关性以及泛化能力，本论文使用设备内准确度 (intra-camera accuracy) 与设备间准确度 (inter-camera accuracy) 两个指标对光源颜色估计的结果进行评估。

为计算设备内准确度这标，在颜色恒常性研究领域中普遍使用的两个基准数据集——ColorChecker RECommended 图像数据集<sup>[186]</sup>以及 NUS-8 Camera 图像数据集<sup>[187]</sup>——的基础上，我们采用三折交叉检验 (3-fold cross validation) 的方法对 ReWNet 模型进行训练与测试。ColorChecker RECommended 图像数据集可视为 Gehler-Shi 数据集<sup>[188]</sup>的改进版本，针对后者中存在的部分光源颜色标签有误的问题，Hemrit 等人对其进行了手动修正。ColorChecker RECommended 数据集共包含 568 张 raw 图像，分别由 Canon 1D 与 Canon 5D 两台相机进行拍摄，覆盖了日常摄影应用中的各种常见场景。NUS-8 Camera 图像数据集则由来自 8 台不同品牌型号相机所拍摄的 1736 张 raw 图像构成，其中包含的场景数量相对于 ColorChecker RECommended 数据集略少一些，但由于每个场景均对应了 8 台相机的实际捕获图像，因此十分适合用于考察算法性能与设备特性间的相关程度。对于 ColorChecker RECommended 与 NUS-8 Camera 数据集中的每台相机，我们对其进行独立的训练与测试，并最终使用各台相机结果间的几何平均值作为性能度量指标。

在评价模型的设备内准确度时，我们直接使用设备相关颜色空间中的 raw 图像作为网络输入，除了进行必要的暗电流扣除外，并未对其进行任何可能改变颜色信息的预处理操作。

在对设备间准确度这一指标进行评价时，除了使用 ColorChecker RECommended 与 NUS-8 Camera 数据集中 10 台相机的图像数据之外，我们还额外加入了 Cube 数据集<sup>[189]</sup>中由 Canon 550D 相机拍摄的 1365 张室外场景的 raw 图像。对于三个基准图像数据库中的 11 台相机，本论文使用留一法交叉检验 (leave-one-out cross validation) 对 ReWNet 模型进

行训练与测试，即在每一轮的评价过程中，使用 10 台相机的图像数据作为训练样本对网络参数进行训练，并使用剩余 1 台相机的图像数据对其进行测试；重复该过程 11 轮并计算其算数平均值，即可得到最终的设备间准确度指标。

由于 11 台不同品牌型号的相机之间存在显著的响应特性差异，因此在将图像输入网络之前，有必要对其进行一定的预处理以确保来自不同相机的图像近似地处于一个共用的颜色空间之中。具体来说，本论文使用如下步骤对图像及光源颜色标签数据进行预处理：

- 1) 扣除暗电流；
- 2) 针对不同的相机，使用 Ddraw 工具包<sup>[168]</sup>提供的  $3 \times 3$  颜色校正矩阵将图像由设备相关颜色空间转换至线性 sRGB 颜色空间<sup>[169]</sup>中；
- 3) 根据  $3 \times 3$  颜色校正矩阵对光源颜色标签进行相应的修正。

为方便起见，下文中将该包含 11 台相机共  $568+1736+1365=3669$  张修正图像与光源颜色标签的数据集称为 MultiCam 数据集\*。

最后，考虑到现有基准数据集中的数据量相对于训练深度网络所需要的海量数据而言仍然有所欠缺（尤其在对设备内准确度进行评价时），我们使用以下数据增强（data augmentation）方法对训练数据进行扩充，以尽可能避免训练过程出现过拟合：

- 从原始分辨率图像中随机裁剪出分辨率在  $256 \times 256$  至  $768 \times 768$  间的子图像（包含标准色卡的图像区域将被自动排除）；
  - 在  $-10^\circ$  至  $+10^\circ$  范围内对子图像进行随机旋转；
  - 以 50% 的概率对图像中的光源颜色进行调整，并保证调整后的光源颜色与真实颜色间的角度差不超过  $5^\circ$ 、 $D_{uv}$  不超过 0.006（光源颜色标签也进行相应的调整）†；
  - 完成数据增强操作后，将图像尺寸重新缩小至  $224 \times 224$  并输入至网络。
- 具体操作细节

本论文使用 TensorFlow 深度学习框架<sup>[191]</sup>作为 ReWNet 的模型搭建平台，并使用 Adam 梯度下降算法<sup>[192]</sup>作为优化器对网络参数进行迭代更新。与[175]、[173]等研究类似，我们使用模型估计出的光源颜色向量（即网络的最终输出）与真实颜色向量间的角度差（参见

\* MultiCam 数据集可于 [https://1drv.ms/f/s!AniPeh\\_FIASDhG7QE6PJiy6s-RIE](https://1drv.ms/f/s!AniPeh_FIASDhG7QE6PJiy6s-RIE) 进行下载。

† 本论文将光源颜色从设备相关颜色空间转换至 CIE 1960 均匀色品标尺图<sup>[161]</sup>后进行  $D_{uv}$  的计算。该约束条件能够确保调整后的光源色品坐标与原光源的色品坐标在等色温线方向上不至于相距过远，否则将与实际拍摄场景中光源颜色分布规律相违背。 $D_{uv} = 0.006$  通常被视作一个“白光”光源所允许的偏离黑体轨迹的最大距离<sup>[190]</sup>。

式(4.38)作为训练阶段的目标函数。训练过程初始阶段所使用的学习率为 $5\times10^{-5}$ ，训练过程中，若检测到目标函数连续6轮内未出现下降，则自动将学习率降低10%。此外，对于多层感知机中的4层全连接层，在每次对目标函数进行反向传播时，其中50%的神经元将被随机选取并忽略（50% dropout<sup>[193]</sup>）以进一步降低过拟合的风险。

同时，我们也对ReWNet模型的空间、时间复杂度进行了具体分析。表4-4中将三种具有不同层级数的ReWNet实例与已有的一些基于深度网络的光源颜色估计模型进行了模型参数数量以及运算效率的定量比较，从中可知，相较于已有的深度网络模型，ReWNet仅需1%~3%的参数规模以及8%~20%的计算量，因此十分适合部署于对模型复杂度与功耗具有较高限制的移动平台之中。特别地，对于具有三层卷积单元的ReWNet实例（3-Hierarchy），其在NVIDIA 1080-Ti GPU上的执行效率约为48ms/张图像，接近于实时运算所需的帧率。

表4-4 基于卷积神经网络的光源颜色估计模型在空间、时间复杂度上的比较（假设输入图像的尺寸均为224×224）。表中的运算次数表示模型在单次正向传递过程中所需要的加法或乘法操作次数

模型名称	参数数量	运算次数
DS-Net <sup>*[178]</sup>	~17.3M	~ $6.0\times10^{10}$
Semantic-CC <sup>[177]</sup>	~13.9M	~ $4.1\times10^9$
FC4 (AlexNet) <sup>[175]</sup>	~3.8M	~ $4.8\times10^9$
Deep Outdoor CC <sup>[176]</sup>	~3.7M	~ $2.3\times10^9$
ReWNet, 1-Hierarchy	7.6K	~ $5.9\times10^7$
ReWNet, 2-Hierarchy	32.7K	~ $1.9\times10^8$
ReWNet, 3-Hierarchy	112.6K	~ $4.7\times10^8$

#### 4.3.2.3 结果与分析

- 设备内准确度

设备内光源颜色估计准确度的评估将对ColorChecker RECommended与NUS-8 Camera数据集中的每台相机独立进行，每个数据集内所有相机评估结果的几何平均值将被作为该

<sup>\*</sup>原生的DS-Net模型仅允许使用44×44尺寸的图像作为输入，因此将图像尺寸增大至224×224将大幅增加其所需的运算次数。

数据集上最终的准确度评价指标。与 4.3.1.5 小节类似，这里同样使用角度差的均值、中位数、三均值、最好 25% 均值以及最差 25% 均值共五种度量方式对 ReWNet 模型的光源颜色估计准确度进行评价。表 4-5 与表 4-6 分别列出了三种 ReWNet 实例在 ColorChecker RERecommended 与 NUS-8 Camera 数据集上与一些典型的光源颜色估计模型的比较结果。需要说明的是，根据 Hemrit 等人的调查研究结果<sup>[186]</sup>，部分算法在对 Gehler-Shi 数据集<sup>[188]</sup>（即 ColorChecker RERecommended 数据集的原始版本）进行评估时有可能使用了不同的光源颜色标签，因此表 4-5 中的部分数据并不适合用于定量比较。

表 4-5 ColorChecker RERecommended 数据集上的设备内 (intra-camera) 光源颜色估计准确度比较

算法	均值	中位数	三均值	最好 25% 均值	最差 25% 均值
White Patch <sup>[165]</sup>	7.55	5.68	6.35	1.45	16.12
Gray World <sup>[166]</sup>	6.36	6.28	6.28	2.33	10.58
1st Order Gray edge <sup>[43]</sup>	5.33	4.52	4.73	1.86	10.03
Shade of Gray <sup>[167]</sup>	4.93	4.01	4.23	1.14	10.20
Beyasian <sup>[57]</sup>	4.82	3.46	3.88	1.26	10.49
Pixels-based Gamut <sup>[194]</sup>	4.20	2.33	2.91	0.50	10.72
Spatio-spectral Statistics <sup>[195]</sup>	3.59	2.96	3.10	0.95	7.61
Bianco CNN <sup>[59]</sup>	2.63	1.98	2.10	0.72	3.90
Cheng et al. (2015) <sup>[58]</sup>	2.42	1.65	1.75	0.38	5.87
FFCC (model Q) <sup>[174]</sup>	2.01	1.13	1.38	0.30	5.14
CCC <sup>[173]</sup>	1.95	1.22	1.38	0.35	4.76
DS-Net <sup>[178]</sup>	1.90	1.12	1.33	0.31	4.84
FC4 (SqueezeNet) <sup>[175]</sup>	1.65	1.18	1.27	0.38	3.78
ReWNet, 1-Hierarchy	2.41	2.02	2.00	0.59	5.10
ReWNet, 2-Hierarchy	2.18	1.73	1.82	0.53	4.70
ReWNet, 3-Hierarchy	1.98	1.38	1.52	0.51	4.52

表 4-6 NUS-8 Camera 数据集上的设备内 (intra-camera) 光源颜色估计准确度比较

算法	均值	中位数	三均值	最好 25% 均值	最差 25% 均值
White Patch <sup>[165]</sup>	10.62	10.58	10.49	1.86	19.45
Pixels-based Gamut <sup>[194]</sup>	7.70	6.71	6.90	2.51	14.05
Gray World <sup>[166]</sup>	4.14	3.20	3.39	0.90	9.00
Beyasian <sup>[57]</sup>	3.67	2.73	2.91	0.82	8.21
Shade of Gray <sup>[167]</sup>	3.40	2.57	2.73	0.77	7.41
1st Order Gray edge <sup>[43]</sup>	3.20	2.22	2.43	0.72	7.36
Spatio-spectral Statistics <sup>[195]</sup>	2.96	2.33	2.47	0.80	6.18
Cheng et al. (2014) <sup>[187]</sup>	2.92	2.04	2.24	0.62	6.61
DS-Net <sup>[178]</sup>	2.24	1.46	1.68	0.48	6.08
CCC <sup>[173]</sup>	2.38	1.48	1.69	0.45	5.85
FC4 (SqueezeNet) <sup>[175]</sup>	2.23	1.57	1.72	0.47	5.15
Cheng et al. (2015) <sup>[58]</sup>	2.18	1.48	1.64	0.46	5.03
FFCC (model Q) <sup>[174]</sup>	2.06	1.39	1.53	0.39	4.80
ReWNet, 1-Hierarchy	2.84	1.92	2.04	0.80	5.82
ReWNet, 2-Hierarchy	2.32	1.64	1.67	0.46	5.44
ReWNet, 3-Hierarchy	2.18	1.59	1.74	0.48	5.35

- 设备间准确度

我们将 MultiCam 数据集中 11 轮留一法交叉检验结果的算数平均值作为设备间光源颜色估计准确度的评价指标，其结果如表 4-7 所示。相比于表 4-5 与表 4-6 中的设备内光源颜色估计结果，表 4-7 中的各项角度差指标均出现了较为显著的劣化（尤其对于基于学习的算法）。经分析我们推测 MultiCam 数据集中不同相机间的响应特性差异是造成这一性能下降的主要原因。在构建 MultiCam 数据集时，虽然尝试了使用颜色校正矩阵将图像响应值从设备相关颜色空间转换至设备无关颜色空间，然而各台相机间的其它一些特性差异，例如响应的线性程度、动态范围大小、传感器的信噪比等，仍然有可能导致转换后的图像无法严格地处于同一样本空间中。训练过程中，这些样本间的差异将不可避免地导致模型的收敛程度出现劣化，从而影响最终的光源颜色估计准确度。

表 4-7 MultiCam 数据集上的设备间 (inter-camera) 光源颜色估计准确度比较

算法	均值	中位数	三均值	最好 25% 均值	最差 25% 均值
Gray World <sup>[166]</sup>	7.85	6.51	6.86	1.27	16.10
Pixels-based Gamut <sup>[194]</sup>	6.12	4.57	4.99	1.32	13.45
Spatio-spectral Statistics <sup>[195]</sup>	5.94	5.01	5.22	1.98	11.29
White Patch <sup>[165]</sup>	5.78	4.37	4.73	1.31	12.64
1 <sup>st</sup> Order Gray edge <sup>[43]</sup>	5.46	4.25	4.52	1.37	11.67
Shade of Gray <sup>[167]</sup>	5.39	4.19	4.48	1.13	11.68
Beyasian <sup>[57]</sup>	5.30	4.10	4.42	1.25	12.14
2 <sup>nd</sup> Order Gray edge <sup>[43]</sup>	5.30	4.01	4.33	1.30	11.50
Bianco CNN <sup>[59]</sup>	3.85	2.82	3.00	1.04	8.50
FFCC (model Q) <sup>[174]</sup>	3.15	2.43	2.60	0.67	6.88
FC4 (SqueezeNet) <sup>[175]</sup>	3.03	2.26	2.35	0.79	6.50
ReWNet, 1-Hierarchy	3.96	3.00	3.29	0.91	8.15
ReWNet, 2-Hierarchy	3.49	2.77	2.89	0.88	7.33
ReWNet, 3-Hierarchy	3.20	2.42	2.58	0.83	7.04

- 可视化展示

图 4-27 对 ReWU 的颜色特征提取机制进行了可视化展示。正如预期的那样，第 0 层的 ReWU（直接与输入 RGB 图像相连接的 ReWU）将输入图像中的中性色区域（见子图 a、f）、镜面反射区域（见子图 g、h）以及肤色（见子图 a~c）、植物的绿色（见子图 d、e、i）等在图像数据集中出现频率较高的记忆色区域赋予了较高的权重；另一方面，对于图像中某些出现频率较低或容易对光源颜色估计造成干扰的颜色，例如暗红色（子图 b 中墙上的油漆）、橘黄色（子图 i 中的辣椒）、藏青色（子图 j 中的汽车漆面）等，ReWU 对其赋予极低的权重，大大弱化了此类像素在网络后端使用全连接层对光源颜色进行回归时所起的作用。

ReWU 所具有的高可解释性允许我们直观地对其工作原理进行理解，其展示出的颜色特征提取机制甚至可以帮助我们在统计类算法的基础上直接对用于像素搜索的约束条件进行人为调节，从而为某些轻量级的自动白平衡算法提供新的设计思路。

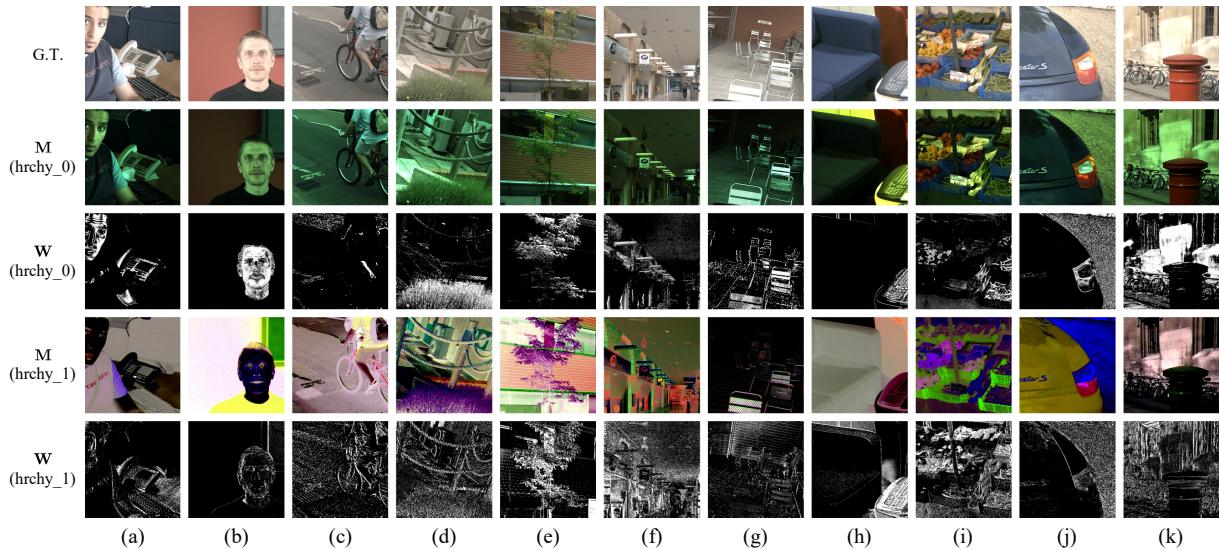


图 4-27 ColorChecker RECommended 数据集上 1-Hierarchy 模型各层级 ReWU 的输入特征图 **M** 与权重分配图 **W**。由上至下：手动白平衡校正结果（已转换至 sRGB 颜色空间）、第 0 层 ReWU（即直接连接输入图像的 ReWU）的输入特征图（即 ReWNet 的输入图像本身）、第 0 层 ReWU 的权重分配图、第 1 层 ReWU 的输入特征图（仅展示 16 个通道中响应值最大的 3 个通道）以及第 1 层 ReWU 的权重分配图

#### 4.4 颜色校正

白平衡校正模块通过对未知光源下拍摄图像的三通道响应值进行独立调节，从而将其中的中性色像素校正至参考光源下的响应值。然而，由于相机光谱灵敏度函数与标准观察者色匹配函数之间的差异，以及未知光源与参考光源光谱功率分布函数之间的差异，白平衡校正模块并无法保证图像中的其它颜色(chromatic colors)也能得到正确的显示。图 4-28 对 X-Rite Classic 标准色卡中各个色块的“真实”响应值以及白平衡校正后的 Nikon D3x 相机响应值进行了比较，其中，前者指的是该色块在 D65 光源下的 XYZ 三刺激值转换至 sRGB 颜色空间后的响应值，而后者则是对 A 光源下预测得到的原始响应值(参见式 (3.63))进行手动白平衡校正后的结果。显然，在完成白平衡校正后，标准色卡中的中性色色块(#19~#24 色块)均已具有正确的颜色外貌，但是其它非中性色色块(#1~#18 色块)与目标颜色之间仍然存在较为显著的颜色差异。

为了解决这一问题，在典型的 ISP 流程中，通常还需要利用颜色校正(color correction)模块将图像中的非中性色像素响应值校正至参考光源下的响应值，以实现严格意义上的颜色恒常性。

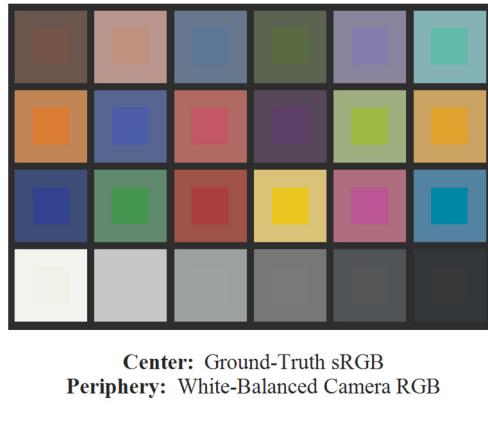


图 4-28 X-Rite Classic 标准色卡的“真实”颜色与白平衡校正后的响应值之间的比较。各色块中心区域的颜色表示其在 D65 光源下的真实响应值 (Ground-Truth)，周围区域的颜色则表示白平衡校正后的 Nikon D3x 相机响应值 (原始响应值在 A 光源下获得)

颜色校正与色度特征化技术 (colorimetric characterization) 无论是在研究目标还是具体的实现方法上均具有很高的相似性。两者的本质区别在于，色度特征化仅负责实现同一颜色信号在不同颜色空间之间的变换，在变换前后场景中的照明光源通常保持不变，因此其补偿的仅仅是成像设备光谱灵敏度函数与标准观察者色匹配函数间的差异；而颜色校正在对颜色空间进行变换的同时，还隐式地对不同照明光源所导致的响应值差异进行了修正，因此其补偿的是“成像设备光谱灵敏度函数与未知光源光谱功率分布的乘积”与“标准观察者色匹配函数与参考光源光谱功率分布的乘积”间的差异。从某种意义上来说，色度特征化可视为一种仅包含单一光源的颜色校正特例。此外，色度特征化模型通常直接以相机的原始响应值作为输入，而颜色校正模块由于需要与白平衡校正模块配合工作，因此其输入往往是经白平衡校正后的响应值。

本节中，对于待标定光源  $U$  下的某个目标样本，我们将其白平衡校正后的相机线性响应值记为  $\mathbf{D}_{\text{wb}}^U$ ，并将该样本在参考光源下的 XYZ 三刺激值记为  $\mathbf{X}^C$ 。颜色校正的本质，实际上是利用一组校正函数  $f_{\text{cc}}$  完成  $\mathbf{D}_{\text{wb}}^U$  至  $\mathbf{X}^C$  的映射，且这一校正函数应对  $M$  个不同的颜色样本同时具备良好的拟合精度，以最小化目标函数

$$\mathcal{L}_{\text{cc}} = \sum_{i=1}^M \mathcal{L}\left[f_{\text{cc}}(\mathbf{D}_{\text{wb},i}^U), \mathbf{X}_i^C\right] \quad (4.56)$$

式中  $\mathcal{L}(\cdot, \cdot)$  表示两组 XYZ 三刺激值之间的误差。

#### 4.4.1 基于根式多项式矩阵的颜色校正算法

由式 (3.66) 可知, 当曝光参数 ( $T$  或  $\mathbf{g}$ ) 发生变化时, 相机原始响应值也将随之变化。因此, 为了确保训练阶段计算得到的颜色校正参数能够适用于不同曝光参数下拍摄的测试图像, 颜色校正模型  $f_{cc}$  应具有一阶齐次 (1-degree homogeneous) 的特性, 即, 当  $y = f_{cc}(x)$  成立时, 应有

$$\alpha y = f_{cc}(\alpha x), \quad \forall \alpha > 0 \quad (4.57)$$

此外, 由于白平衡校正模块在颜色校正之前已经对图像进行了三通道增益调节, 使得图像中的中性色像素具有正确的响应值, 因此颜色校正模型还应该具有中性色恒定性, 即, 对于经白平衡校正后的光源  $U$  下的任意一组中性色响应值, 其在颜色校正后也应对应参考光源下的中性色三刺激值:

$$k\mathbf{X}_n^C = f_{cc}\left(k\mathbf{D}_{n,wb}^U\right), \quad \forall k \in [0, 1] \quad (4.58)$$

其中  $\mathbf{D}_{n,wb}^U = [1, 1, 1]$  表示白平衡校正后的白点响应值,  $\mathbf{X}_n^C$  表示参考白点的色度三刺激值, 若使用 D65 作为参考光源, 则有  $\mathbf{X}_n^C = [0.95, 1, 1.09]$ <sup>[196]</sup>。当满足一阶齐次特性(即式 (4.57))时, 式 (4.58) 可等价于

$$\mathbf{X}_n^C = f_{cc}\left(\mathbf{D}_{n,wb}^U\right) \quad (4.59)$$

在综合考虑了颜色校正精度、泛化能力以及模型复杂度后, 本论文选择使用二次项的齐次根式多项式模型 (2nd degree root-polynomial color correction model)<sup>[63]</sup> 作为颜色校正模块中所使用的映射模型 (以下简称二次项 RPCC 模型), 同时使用 CIEDE2000 色差作为目标函数, 以提高颜色校正结果与人眼视觉系统实际感知的颜色差异之间的一致性。

对于输入图像中每个像素对应的  $1 \times 3$  原始响应值向量  $\mathbf{D} = [D_r, D_g, D_b]$ , 二次项 RPCC 模型首先对其进行根式多项式扩展, 得到  $1 \times 6$  的根式多项式行向量:

$$\tilde{\mathbf{D}} = [D_r, D_g, D_b, \sqrt{D_r D_b}, \sqrt{D_r D_b}, \sqrt{D_g D_b}] \quad (4.60)$$

接着, 二次项 RPCC 模型将该根式多项式向量右乘一个  $6 \times 3$  的颜色校正矩阵, 以得到颜色校正后参考光源下对应的 XYZ 估计值:

$$\begin{aligned} \mathbf{D}_{cc} &= [\hat{X}, \hat{Y}, \hat{Z}] = \tilde{\mathbf{D}} \cdot \mathbf{M} \\ &= [D_r, D_g, D_b, \sqrt{D_r D_b}, \sqrt{D_r D_b}, \sqrt{D_g D_b}] \cdot \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ \vdots & \vdots & \vdots \\ m_{61} & m_{62} & m_{63} \end{bmatrix} \end{aligned} \quad (4.61)$$

当光源发生变化时，相机原始响应值也将随之变化，因此为保证颜色校正结果的准确性，通常需要针对不同的光源分别对二次项 RCPP 颜色校正矩阵进行计算与优化。本论文使用 D65、A、D50、CWF、TL84 等五种典型光源作为标定光源对颜色校正模型进行标定，并将 D65 作为参考光源<sup>\*</sup>。本论文基于原始响应值预测模型对二次项 RCPP 颜色校正矩阵进行标定，其具体原理及实现过程请参阅第 5.2.1 小节。

为了对颜色校正模型的泛化性能进行评估，本论文使用 X-Rite Digital SG 色卡中的 140 个色块作为训练样本对颜色校正矩阵  $\mathbf{M}(U)$  进行计算，并在同一光源下使用 X-Rite Classic 色卡中的 24 个色块作为测试样本对其校正精度进行检验。表 4-8 列出了不同光源下两台相机在测试样本上的颜色校正精度，其中中性色平均色差表示对 X-Rite Classic 色卡中的 #19~#24 色块（即图 4-28 中最下方的 6 个中性色色块）计算色差均值后的结果。由表 4-8 中的色差数据可以看出，经非线性优化后的颜色校正矩阵对于不同的光源均具有良好的颜色校正精度以及泛化能力（作为参照，工业界中通常将  $\Delta E_{ab}^* = 2.3$  作为恰可察觉色差<sup>[139, 197]</sup>）；同时，得益于白点约束的引入，测试样本中中性色色块的色准在颜色校正后也得到了很好的保证。图 4-29 中展示了不同光源下使用二次项 RPCC 颜色校正矩阵对 Nikon D3x 拍摄的 X-Rite Classic 色卡图像进行颜色校正前后各测试色块的颜色外貌对比。

表 4-8 Nikon D3x (N) 与 Sony ILCE7 (S) 在各光源下的颜色校正精度

光源 $U$	平均 $\Delta E_{00}$ 色差		最大 $\Delta E_{00}$ 色差		中性色平均 $\Delta E_{00}$ 色差		平均 $\Delta E_{ab}^*$ 色差		最大 $\Delta E_{ab}^*$ 色差		中性色平均 $\Delta E_{ab}^*$ 色差	
	N	S	N	S	N	S	N	S	N	S	N	S
D65	1.07	0.98	4.31	3.35	1.26	1.39	1.57	1.39	5.21	4.09	1.24	4.36
A	1.68	1.56	9.48	9.05	0.99	1.02	2.85	2.65	12.03	11.42	0.97	1.00
D50	1.19	1.09	5.64	4.84	1.18	1.28	1.83	1.60	6.87	5.88	1.17	1.26
CWF	1.07	1.06	6.16	5.54	1.18	1.28	1.82	1.78	7.57	6.74	1.12	1.20
TL84	1.36	1.33	6.37	5.88	1.13	1.20	2.27	2.19	9.19	8.74	1.07	1.14

\* 当参考光源与待校正光源同为 D65 时，颜色校正的过程与色度特征化等价（参见第 3.3.3.2 小节）。

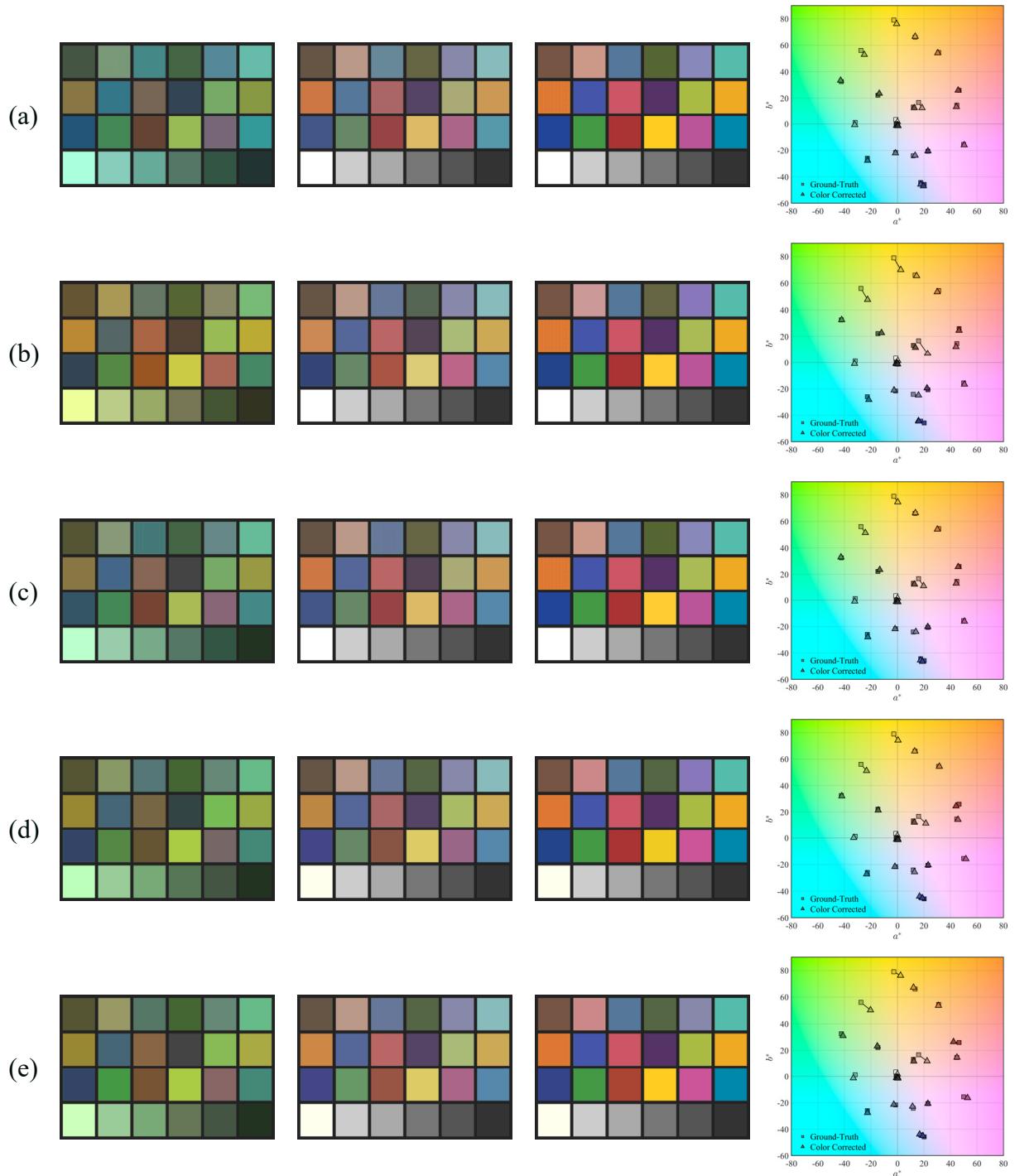


图 4-29 Nikon D3x 所拍摄的 X-Rite Classic 色卡在 (a) D65、(b) A、(c) D50、(d) CWF 以及 (e) TL84 光源下颜色校正前后的颜色外貌对比。由左至右：白平衡校正之前的各色块原始响应值、白平衡校正后的响应值、颜色校正后的响应值以及 CIE1976  $a^*b^*$  色度平面的误差展示（色度平面中每个方块代表一个测试色块的参考色品坐标，三角形代表颜色校正后的色品坐标）

#### 4.4.1.1 颜色校正流程

针对单个标定光源  $U$  完成颜色校正矩阵的标定后，可将该光源对应的白平衡增益系

数  $\mathcal{G}_r^U$ 、 $\mathcal{G}_b^U$  以及颜色校正矩阵  $\mathbf{M}(U)$  以“键-值”对的形式进行存储。在对所有标定光源完成这一流程后，即得到了颜色校正所需的全部标定参数。

对于输入至颜色校正模块中的任意白平衡校正后的图像  $\mathcal{I}_{\text{wb}}$  及其白平衡增益系数  $\mathcal{G}_r^T$ 、 $\mathcal{G}_b^T$ （由自动白平衡校正模块进行估计或由用户手动指定），本论文采取以下流程对其进行颜色校正：

1) 利用式 (3-73) 对输入图像  $\mathcal{I}_{\text{wb}}$  进行非线性逆变换，得到线性图像  $\mathcal{I}_{\text{wb},0}$ 。

2) 利用式 (3-73) 对输入图像白平衡增益系数的倒数  $1/\mathcal{G}_r^T$ 、 $1/\mathcal{G}_b^T$  进行非线性逆变换并再次计算倒数，得到线性增益系数  $\mathcal{G}_r^{\mathcal{I}_0}$ 、 $\mathcal{G}_b^{\mathcal{I}_0}$

3) 利用公式

$$\begin{cases} D_{n,r}^U = \mathcal{G}_r^U / \mathcal{G}_r^C \\ D_{n,g}^U = 1 \\ D_{n,b}^U = \mathcal{G}_b^U / \mathcal{G}_b^C \end{cases} \quad (4.62)$$

对每个标定光源的相机响应值进行计算，其中  $\mathcal{G}_r^C$ 、 $\mathcal{G}_b^C$  表示参考光源对应的白平衡增益系数，下标“ $n$ ”表示该响应值是针对中性色物体表面计算得到的（与光源自身的响应值等价）。

4) 利用经验公式<sup>[169]</sup>

$$\begin{bmatrix} \hat{X}_n^U \\ \hat{Y}_n^U \\ \hat{Z}_n^U \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.2127 & 0.0193 \\ 0.3576 & 0.7152 & 0.1192 \\ 0.1804 & 0.0722 & 0.9503 \end{bmatrix} \begin{bmatrix} D_{n,r}^U \\ D_{n,g}^U \\ D_{n,b}^U \end{bmatrix} \quad (4.63)$$

对各个标定光源的 XYZ 三刺激值进行估算，并将其转换至 CIE1976 UCS 色品坐标以保证颜色均匀性<sup>[67]</sup>：

$$\begin{cases} \hat{u}_n^U = \frac{4\hat{X}_n^U}{\hat{X}_n^U + 15\hat{Y}_n^U + 3\hat{Z}_n^U} \\ \hat{v}_n^U = \frac{9\hat{Y}_n^U}{\hat{X}_n^U + 15\hat{Y}_n^U + 3\hat{Z}_n^U} \end{cases} \quad (4.64)$$

5) 对于输入图像的线性增益系数  $\mathcal{G}_r^{\mathcal{I}_0}$ 、 $\mathcal{G}_b^{\mathcal{I}_0}$ ，同样使用式 (4.62)~(4.64) 计算其 CIE1976 UCS 色品坐标  $\hat{u}_n^{\mathcal{I}_0}$ 、 $\hat{v}_n^{\mathcal{I}_0}$ 。

6) 在  $(u, v)$  平面内计算各个标定光源的色品坐标与输入图像光源色品坐标的欧氏距离，并选取距离最小的标定光源对应的颜色校正矩阵作为输入图像的颜色校正矩阵：

$$\begin{aligned} \mathbf{M}(\mathcal{I}_{wb,0}) &= \mathbf{M}(U_i) \\ i &= \arg \min \sqrt{\left(\hat{u}_n^{U_i} - \hat{u}_n^{\mathcal{I}_0}\right)^2 + \left(\hat{v}_n^{U_i} - \hat{v}_n^{\mathcal{I}_0}\right)^2} \end{aligned} \quad (4.65)$$

7) 使用颜色校正矩阵  $\mathbf{M}(\mathcal{I}_{wb,0})$  对线性图像  $\mathcal{I}_{wb,0}$  进行颜色校正, 即, 对于  $\mathcal{I}_{wb,0}$  中每一个像素的响应值  $\mathbf{D}_0 = [D_{r,0}, D_{g,0}, D_{b,0}]$ , 利用

$$\mathbf{D}_{cc} = \tilde{\mathbf{D}}_0 \cdot \mathbf{M}(\mathcal{I}) \quad (4.66)$$

计算其颜色校正后的响应值, 其中  $\tilde{\mathbf{D}}_0$  表示  $\mathbf{D}_0$  的二次项根式多项式扩展 (参见式(4.60))。

图 4-30 中展示了使用以上方法对 Nikon D3x ColorChecker 图像数据集中部分测试图像进行颜色校正前后的颜色外貌比较。

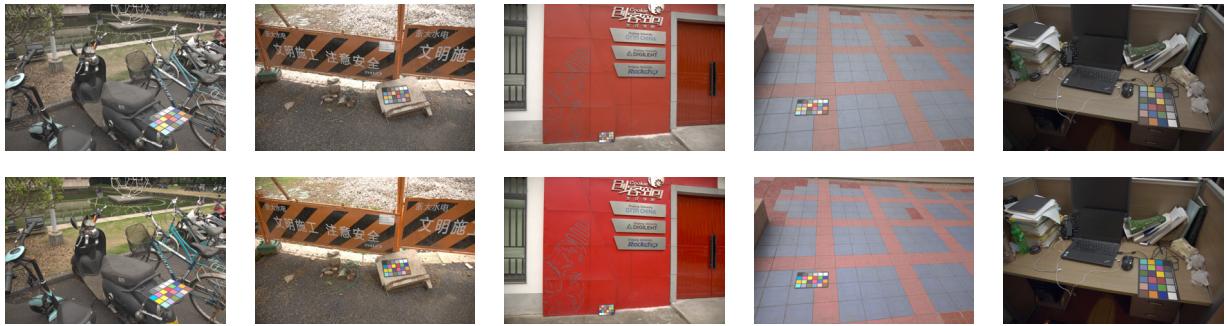


图 4-30 部分测试图像在颜色校正前 (上) 后 (下) 的颜色外貌比较\*

需要说明的是, 虽然式 (4.63) 中的颜色空间变换矩阵仅适用于使用线性 sRGB 驱动值作为输入的情况, 然而对于所有标定光源而言, 其响应值向量  $[D_{n,r}^U, D_{n,g}^U, D_{n,b}^U]$  与线性 sRGB 颜色空间中的白点坐标  $[1,1,1]$  之间的差异并不显著, 因此可以近似地将白平衡后的光源颜色响应值视作线性 sRGB 驱动值。

#### 4.4.1.2 结果与分析

为了考察颜色校正算法在真实图像上的校正性能以及泛化能力, 本论文使用 Nikon D3x ColorChecker 数据集 (详见 4.3.1.4 小节) 对最终输出图像的颜色复现精度进行定量评估。

对于 Nikon D3x ColorChecker 数据集中的每张测试图像, 我们以光源真实的 RGB 响应值为参考, 手动指定其白平衡增益系数  $\mathcal{G}_r^T$ 、 $\mathcal{G}_b^T$ , 并使用上一段落中的操作流程对其进

\* 图 4-30 中展示的均为 ISP 输出图像。除颜色校正模块之外其余各模块均为手动校正。

行颜色校正。完成颜色校正后，从测试图像中标准色卡所在区域手动提取出各色块的响应值，并与 X-Rite Classic 色卡在参考光源下的理想响应值进行色差分析，以此作为颜色复现精度的评价指标。由于各张测试图像的拍摄光环境与曝光参数均存在差异，因此在进行色差分析之前，允许对每张测试图像乘上一个全局的响应值调制常数  $s$  以模拟拍摄时曝光参数的调节。实际操作时，我们以最小化各色块的 CIEDE2000 色差的均值作为计算该响应值调制系数的依据，即

$$s = \arg \min \frac{1}{24} \sum_{i=1}^{24} \Delta E_{00}(s \mathbf{D}_{cc,i}, \mathbf{X}_i^C) \quad (4.67)$$

式中  $\mathbf{D}_{cc,i}$  表示对 X-Rite Classic 色卡中第  $i$  个色块进行颜色校正后的响应值（参见式（4.66））， $\mathbf{X}_i^C$  表示其在参考光源下真实的 XYZ 三刺激值。图 4-31 中展示了使用全局响应值调制系数对其中一张测试图像进行曝光调节前后的图像外貌对比。



图 4-31 使用全局响应值调制系数对测试图像进行曝光调节前（左）后（右）的图像外貌对比\*。本例中  $s = 1.323$

对于每张测试图像，我们以 24 个色块的平均色差作为其颜色复现精度的评价指标。与第 4.3.1.5 小节中光源颜色角度差的分析方法类似，这里使用平均色差的均值、中位数、三均值、最好 25% 均值以及最差 25% 均值共五种度量方法对 Nikon D3x ColorChecker 数据集中的所有测试图像进行分析，其结果如表 4-9 所示。

需要说明的是，使用实际图像中的色卡响应值对颜色复现精度进行评估时，除了颜色校正算法本身的性能将对结果产生影响之外，其余的一些不可控因素（例如拍摄场景中光源的非均匀性、部分图像中色卡区域存在阴影、少数图像存在过曝现象等）均有可能导致最终的颜色复现精度出现下降，这也是为什么实验室环境下测试数据集上的颜色复现精度

\* 图 4-31 中所展示的均为 ISP 输出图像。

(参见表 4-8) 要显著高于实际图像的原因所在。

表 4-9 Nikon D3x ColorChecker 数据集上的颜色复现精度

色差公式	均值	中位数	三均值	最好 25% 均值	最差 25% 均值
$\Delta E_{00}$	3.00	2.57	2.67	1.88	4.80
$\Delta E_{ab}^*$	5.14	4.51	4.65	3.18	8.24

#### 4.4.2 物体色同色异谱失配与颜色失真

物体色同色异谱失配 (object-color metamer mismatching, 下文中在不引起歧义的情况下简称为同色异谱) 指的是在某一光源下具有相同颜色外貌的两个物体表面, 在另一种光源下却呈现出不同的颜色。

为了对颜色校正过程中因同色异谱失配引起的颜色失真进行定量评估, 本小节在高光谱图像数据集上对数码相机成像过程中的同色异谱现象进行仿真, 通过逐像素搜索物体色同色异谱集合的方法生成一系列具有不同程度同色异谱劣化的图像, 并将其与参考图像 (即参考光源下人眼直接观察到的图像) 间的图像质量评价 (image quality assessment) 结果作为其颜色失真程度的评价指标。

暂不考虑响应值构成模型中的非线性与各常数项, 则联立式 (2.9) 与 (3.63) 可得到光谱反射比函数为  $\rho(\lambda)$  的物体表面对应的原始响应值向量  $\mathbf{D}_\rho$ :

$$\mathbf{D}_\rho = \begin{bmatrix} D_r \\ D_g \\ D_b \end{bmatrix}^\top = \begin{bmatrix} \int \rho(\lambda) P(\lambda) s_r(\lambda) \\ \int \rho(\lambda) P(\lambda) s_g(\lambda) \\ \int \rho(\lambda) P(\lambda) s_b(\lambda) \end{bmatrix}^\top \quad (4.68)$$

式中  $P(\lambda)$  表示光源的光谱功率分布函数,  $s_k(\lambda), k \in \{r, g, b\}$  表示数码相机  $k$  通道的光谱灵敏度函数。

由前文可知, 颜色校正模块的目标总是将未知光源下拍摄的图像校正至参考光源下的颜色外貌, 因此针对这一应用背景, 本论文对数码相机颜色复现过程中的物体色同色异谱现象进行如下定义: 对于在未知光源  $U$  下具有相同原始响应值的一对物体表面  $\rho_1$ 、 $\rho_2$ , 若在参考光源  $C$  下具有不同的原始响应值, 则认为  $\rho_1$  与  $\rho_2$  在光源  $U$  下同色异谱。特别地, 对于某一目标光谱  $\rho_0$ , 将光源  $U$  下所有与其同色异谱的光谱反射比构成的集合称为  $\rho_0$  在

$U$  下的物体色同色异谱集合 (object-color metamer set, 下文中简称为 OCMS)，并使用符号  $\mathcal{O}^U(\rho_0)$  进行表示，即

$$\mathcal{O}^U(\rho_0) \triangleq \left\{ \rho \mid \mathbf{D}_\rho^U = \mathbf{D}_{\rho_0}^U \right\} \quad (4.69)$$

本小节中仍然使用 D65 作为参考光源。

#### 4.4.2.1 同色异谱劣化图像的仿真

为考察因同色异谱引起的颜色失真对于最终图像质量的影响，本论文使用同色异谱劣化图像 (metamer-degraded image) 与参考图像  $\mathcal{I}^*$  间的全参考图像质量评价 (full-reference image quality assessment, FR-IQA，以下简称为 IQA) 分数作为颜色失真程度的量化评价指标。为了模拟不同光源下的拍摄的 raw 图像，本论文使用高光谱图像 (hyperspectral image) 作为测试平台<sup>\*</sup>，并利用原始响应值预测模型对其中的每个像素在不同光源下的响应值进行仿真。同时，利用式 (2.12)，可得到高光谱图像中的每个像素在参考光源  $C$  下的 XYZ 三刺激值，在将每个像素的三刺激值转换为 sRGB 驱动值后，即可获得与人眼视觉系统具有一致颜色感知的参考图像  $\mathcal{I}^*$ 。

对于高光谱图像中光谱反射比为  $\rho_0(\lambda)$  的某个像素，我们在一个预先定义的光谱反射比数据库  $\mathcal{R}$  中对其 OCMS 进行搜索<sup>†</sup>。考虑到现实中能够获取的光谱反射比数据量总是有限的，因此为了保证最终获得的  $\mathcal{O}^U(\rho_0)$  中含有足够数量的光谱反射比数据，需要适当放宽式 (4.69) 中对于 OCMS 的定义。正式地，本论文采用如下判据对目标光谱  $\rho_0$  在  $U$  下的物体色同色异谱集合进行定义：

$$\mathcal{O}^U(\rho_0) \triangleq \left\{ \rho \in \mathcal{R} \mid \mathcal{L}\left(\mathbf{D}_\rho^U, \mathbf{D}_{\rho_0}^U\right) \leq \epsilon \right\} \quad (4.70)$$

其中  $\mathcal{L}(\cdot, \cdot)$  表示两组原始响应值间的误差评价指标， $\epsilon$  表示一个预先指定的误差阈值。考虑到用户总是在 ISP 最终的输出图像中对拍摄物体的颜色外貌进行评价，因此本论文使用颜色校正后  $\mathbf{D}_\rho^U$  与  $\mathbf{D}_{\rho_0}^U$  之间的色差作为其误差评价指标，此时，式 (4.70) 可改写为

\* 不同于传统的三通道图像，高光谱图像中的每个像素完整地记录了拍摄场景中该位置处对应物体表面的光谱反射比信息，因此可被视为一种不受照明条件影响的对被摄物体真实反射特性进行记录的存储媒介。

† 关于光谱反射比数据库的更多细节将在下文中介绍。

$$\mathcal{O}^U(\rho_0) \triangleq \left\{ \rho \in \mathcal{R} \mid \Delta E_{ab}^* \left[ f_{cc}^U(\mathbf{D}_\rho^U), f_{cc}^U(\mathbf{D}_{\rho_0}^U) \right] \leq \epsilon \right\} \quad (4.71)$$

式中  $f_{cc}^U(\mathbf{D})$  表示使用  $U$  光源对应的颜色校正矩阵对  $\mathbf{D}$  进行二次项齐次根式多项式颜色校正（详见第 4.4.1.1 小节），同时，我们使用 CIELAB 色差的恰可察觉差作为计算 OCMS 的色差阈值<sup>[139, 197]</sup>，即令  $\epsilon = 2.3$ 。

由于物体色同色异谱失配现象的存在，集合  $\mathcal{O}^U(\rho_0)$  中必然存在某些光谱反射比函数，其在参考光源  $C$  下与  $\rho_0(\lambda)$  的色差不再满足小等于  $\epsilon$  这一条件。对于此类物体表面，其在参考光源下本该具有与  $\rho_0(\lambda)$  不同的颜色外貌，然而当我们使用相机在光源  $U$  下对其拍摄时，这一颜色差异并无法在输出图像上得到体现，因此可认为此时跨光源图像复现的过程中存在一定程度的颜色失真。

为了对具有不同程度颜色失真的图像进行质量评价，本论文为每一幅高光谱图像生成了 20 幅不同劣化程度的同色异谱劣化图像，分别对应于所有可能的同色异谱劣化图像中的第 0、第 5、……以及第 95 百分位数（percentile）。位于第  $p$  个百分位数的同色异谱劣化图像  $\mathcal{I}_{ocm}^{p\%}$  的生成过程简述如下：

- 1) 对于高光谱图像中位于  $x$  处光谱反射比为  $\rho_0(\lambda)$  的像素，使用式 (4.71) 对其 OCMS  $\mathcal{O}^U(\rho_0)$  进行搜索。
- 2) 对于集合  $\mathcal{O}^U(\rho_0)$  中的所有元素，逐一计算其在参考光源下经颜色校正后的响应值  $f_{cc}^C(\mathbf{D}^C)$  与  $\rho_0(\lambda)$  在参考光源下 XYZ 三刺激值之间的 CIELAB 色差：

$$\begin{cases} \Delta_1 = \Delta E_{ab}^* \left[ f_{cc}^C \left( \mathbf{D}_{\rho_1}^C \right), \mathbf{X}_{\rho_0}^C \right] \\ \vdots \\ \Delta_N = \Delta E_{ab}^* \left[ f_{cc}^C \left( \mathbf{D}_{\rho_N}^C \right), \mathbf{X}_{\rho_0}^C \right] \end{cases} \quad (4.72)$$

其中  $N = |\mathcal{O}^U(\rho_0)|$  表示集合  $\mathcal{O}^U(\rho_0)$  中的元素个数。

- 3) 将步骤 2 中的  $N$  组色差结果由小到大进行排序，并将色差位于第  $p$  个百分位数的光谱反射比函数  $\rho_{p\%}(\lambda)$  在参考光源  $C$  下经颜色校正后的响应值赋予  $\mathcal{I}_{ocm}^{p\%}$  中  $x$  处的像素：

$$\mathcal{I}_{ocm}^{p\%}(x) = f_{cc}^C \left( \mathbf{D}_{\rho_{p\%}}^C \right) \quad (4.73)$$

- 4) 对高光谱图像中所有像素重复以上步骤，得到最终的同色异谱劣化图像  $\mathcal{I}_{ocm}^{p\%}$ 。
- 5) 将劣化图像  $\mathcal{I}_{ocm}^{p\%}$  由 XYZ 颜色空间转换至 sRGB 颜色空间，以便于 IQA 模型进行图像质量

评价。

图 4-32 中展示了完整的 OCMS 搜索步骤以及不同劣化程度的同色异谱劣化图像的仿真流程。需要说明的是，在步骤 3 中，我们使用“第 0 百分位数”的说法来表示目标光谱反射比函数本身，即  $\rho_{0\%}(\lambda) = \rho_0(\lambda)$ 。此时， $\mathcal{I}_{\text{ocm}}^{p\%}$  实际上表示的是一幅不存在同色异谱失配的理想图像，其与参考图像  $\mathcal{I}^*$  间的误差将完全来自于颜色校正模型  $f_{\text{cc}}^C$  自身的校正误差。

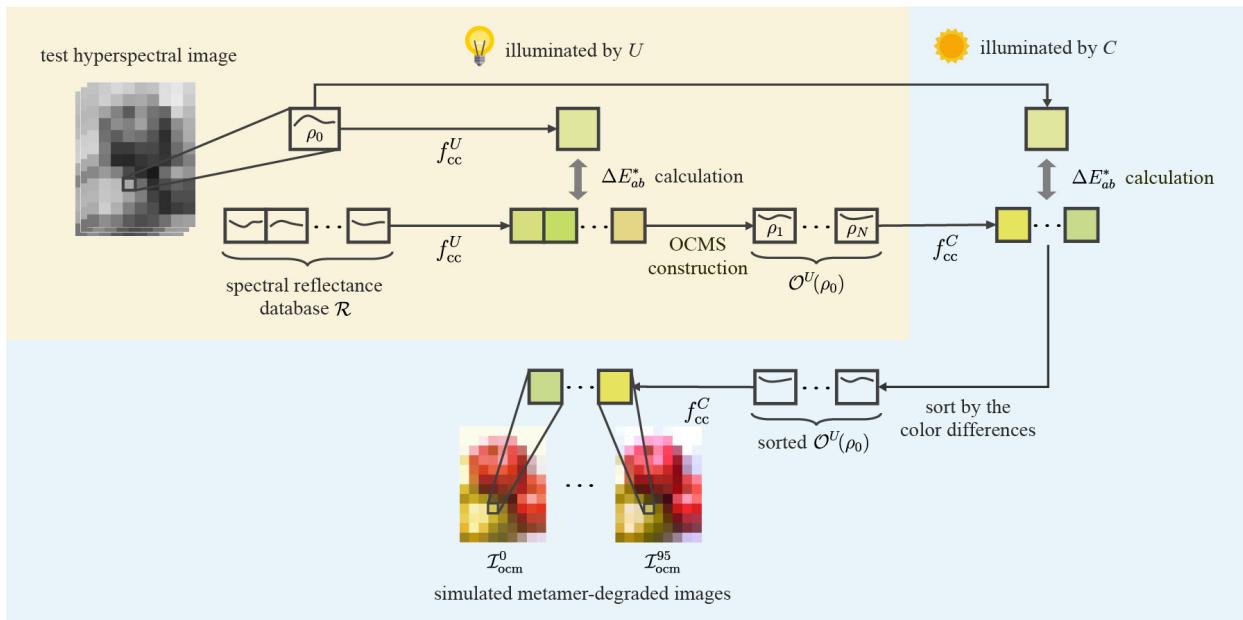


图 4-32 光源  $U$  下不同劣化程度的同色异谱劣化图像仿真流程

上文通过对各个像素的 OCMS 进行独立搜索从而构建出最终的同色异谱劣化图像。然而，实际拍摄场景中相邻的物体表面往往具有较高的相似性，因此在对同色异谱劣化图像进行仿真时，有必要加入一定的“平滑”操作以避免最终的图像中存在明显的纹理走样与瑕疵。基于这一考虑，对于具有光谱反射比函数  $\rho_0(\lambda)$  的目标像素，我们在以该像素为中心的  $3 \times 3$  窗口内对每个像素的 OCMS 进行计算，并使用其并集作为该像素对应的改进版 OCMS：

$$\tilde{\mathcal{O}}^U(\rho_0) \triangleq \bigcup_{\rho_x \in \mathcal{W}(\rho_0)} w_x \mathcal{O}^U(\rho_x) \quad (4.74)$$

式中  $\mathcal{W}(\rho_0)$  表示以  $\rho_0$  为中心的  $3 \times 3$  窗口内各光谱反射比函数构成的集合， $w_x$  表示该窗口内像素  $x$  对应的布尔权重系数，有

$$w_x = \begin{cases} 1 & \text{if } g(\rho_0, \rho_x) \geq \epsilon_{\text{sim}} \\ 0 & \text{otherwise} \end{cases} \quad (4.75)$$

式中  $\epsilon_{\text{sim}}$  表示相似性阈值，本节中令  $\epsilon_{\text{sim}} = 0.995$ ； $g(\cdot, \cdot)$  用于计算两组光谱反射比函数之间的相似性<sup>[198]</sup>，有

$$g(\rho_m, \rho_n) = \frac{\sum_{\lambda} [(\rho_m(\lambda) - \bar{\rho}_m) \cdot (\rho_n(\lambda) - \bar{\rho}_n)]}{\sqrt{\sum_{\lambda} (\rho_m(\lambda) - \bar{\rho}_m)^2 \cdot \sum_{\lambda} (\rho_n(\lambda) - \bar{\rho}_n)^2}} \quad (4.76)$$

式中  $\bar{\rho}_m$  与  $\bar{\rho}_n$  分别表示  $\rho_m$  与  $\rho_n$  在不同波长处的均值。

图 4-33 对比了使用两种 OCMS 搜索方法仿真得到的同色异谱劣化图像的颜色外貌。为表述方便起见，下文仍然使用“物体色同色异谱集合（OCMS）”对式 (4.74) 中的光谱反射比并集进行描述，若无特殊说明，则 OCMS 均特指  $\tilde{\mathcal{O}}^U(\rho_0)$  而非  $\mathcal{O}^U(\rho_0)$ 。



图 4-33 使用  $\mathcal{O}^U(\rho_0)$  (左) 以及  $\tilde{\mathcal{O}}^U(\rho_0)$  (右) 仿真得到的 50% 同色异谱劣化图像  $\mathcal{I}_{\text{ocm}}^{50\%}$ 。该图像使用 F8 荧光灯光源作为测试光源  $U$ ，并以 Nikon D50 作为测试相机

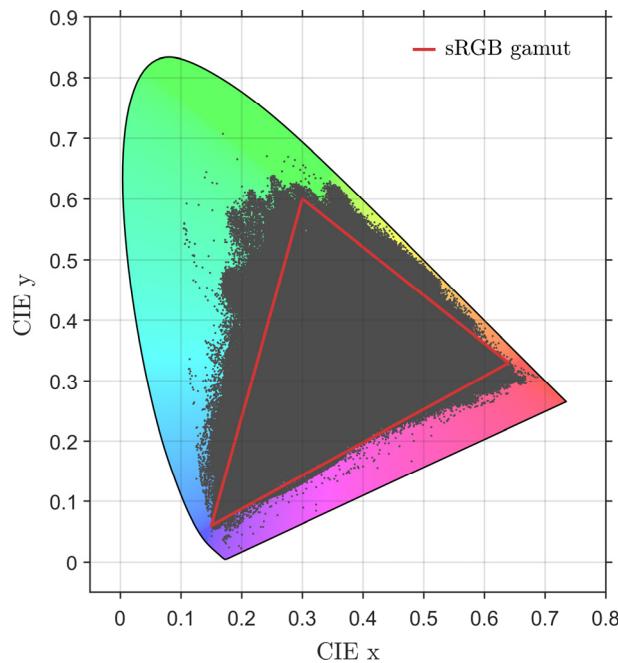
#### 4.4.2.2 实验设置

- 光谱反射比数据库

为了尽可能真实地模拟实际拍摄场景中的同色异谱失配现象，我们从公开的实测光谱数据集<sup>[72, 199-201]</sup>以及高光谱图像（见表 4-10）中对光谱反射比数据进行收集。在对 43,714,192 组原始数据中的重复样本以及存在明显异常的样本进行剔除后，共得到 7,326,497 组有效的光谱反射比数据，其在 CIE1931 xy 色度平面内的色品坐标分布如图 4-34 所示。这 7,326,497 组光谱反射比数据覆盖了实际拍摄场景中大多数的真实物体表面，为 OCMS 搜索提供了有效依据。

表 4-10 高光谱图像数据来源

数据库名称	图像内容	图像数量
Foster <sup>[202]</sup>	植物、建筑、纸制品等	19
Yasuma <sup>[203]</sup>	多种类	32
UEF <sup>[204]</sup>	皮肤、塑料制品等	12
Skauli <sup>[205]</sup>	面部皮肤、植物、建筑等	33
Moan <sup>[206]</sup>	多种类	9
Hordley <sup>[207]</sup>	印刷材料、纸制品等	21
Parraga <sup>[208]</sup>	植物	27
Eckhard <sup>[209]</sup>	户外风景、建筑等	14

图 4-34 光谱反射比数据库  $\mathcal{R}$  中 7,326,497 个样本在 CIE1931  $xy$  色度平面内的色品坐标分布（使用 D65 作为照明光源）

- 测试光源

本实验共选取了 8 种常见的光源对不同照明条件下的成像过程进行模拟，如表 4-11 所示。同时，我们还额外加入了重组日光光谱以便将物体色同色异谱失配与光源色同色异谱失配现象 (illuminant metamer mismatching) 进行比较，即，对于每一种测试光源 (A、D50、D100 以及 TL84 除外，因为重组日光的相关色温必须高于 4000K)，通过对不同基

函数的系数进行调节从而生成与目标光源具有相同相关色温的重组日光光谱<sup>[210]</sup>。图 4-35 中绘制出了 8 种测试光源以及对应的重组日光光源的相对光谱功率分布函数。

表 4-11 八组测试光源及其相关色温

序号	光源 (简称)	相关色温
1	A	2864K
2	D50	5007K
3	D100	9989K
4	Cool White Fluorescent (CWF)	4230K
5	D50 fluorescent simulator (F8)	5000K
6	TL84	3980K
7	D50 LED simulator (LED-D50)	4894K
8	iPhone 7 True Tone flash (Flash)	5071K

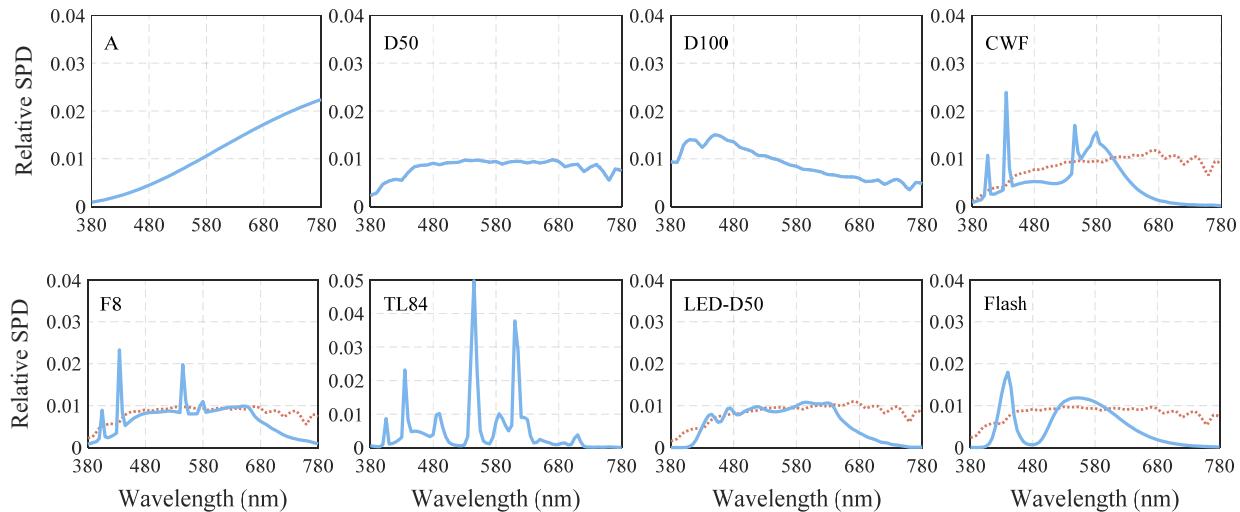


图 4-35 八种测试光源 (蓝色实线) 与相同相关色温的重组日光光源 (红色点线) 的相对光谱功率分布函数 (所有光源均归一化至  $Y=1$ )

- 高光谱测试图像

本实验使用表 4-10 中列出的 8 个数据集中的高光谱图像作为测试数据来源。这些测试图像的空间分辨率由  $256 \times 256$  至  $1344 \times 1024$  不等, 为提高仿真的效率, 对于像素数大于 100 万的测试图像, 本论文采用隔行隔列扫描的方式将其下采样至原始分辨率的四分之一,

同时确保其光谱反射比数据不受影响。

对于高光谱图像中的某个像素，当且仅当其在所有测试光源（包括同色异谱 LED 与重组日光光源）下 OCMS 中的样本数均不小于 100 时我们认为其统计结果具有一定的代表性，并将满足该条件的像素称为同色异谱有效像素（metamer-valid pixel）。为了保证仿真结果的可靠性，我们从图像数据集中剔除掉那些同色异谱有效像素个数与总像素数之比低于 95% 的高光谱图像，最终共获得 136 张有效的测试图像，覆盖了日常拍摄场景中的植物、皮肤、印刷品、纺织品、塑料、建筑等常见物体。

- 测试相机

成像设备的光谱响应特性将对同色异谱导致的颜色失真程度产生直接的影响。一般来说，成像设备的光谱灵敏度函数与人眼视觉系统色匹配函数之间的相似性越低，其颜色失真程度也就越显著。为了对同色异谱造成颜色失真在不同设备间的分布规律进行探讨，我们从 40 台候选机型<sup>[23]</sup>中选取了 Canon 60D 单反相机、Nikon D80 单反相机以及 PointGrey Grasshopper 50S5C 工业相机三台具有代表性的相机进行测试——在所有 40 台机型中，这三台相机分别对应了最高（96.2）、中位数（93.1）以及最低（79.6）的相机灵敏度同色异谱指数（digital still camera sensitivity metamerism index, DSCSMI）\*。图 4-36 中绘制了三台测试相机各自的光谱灵敏度函数。

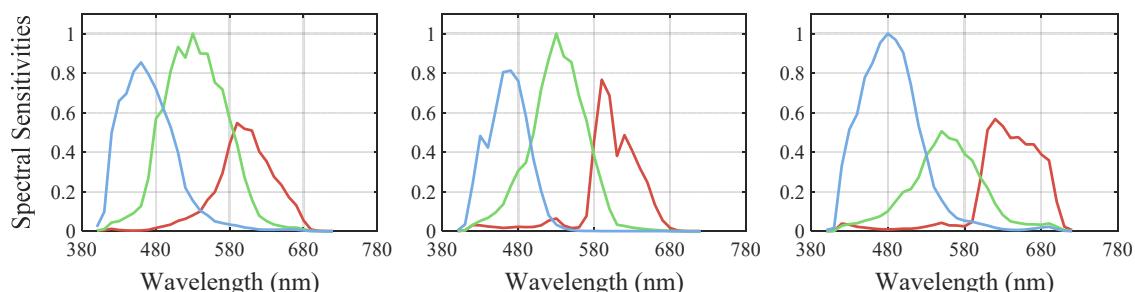


图 4-36 三台测试相机的归一化光谱灵敏度函数。由左至右：Canon 60D 单反相机、Nikon D80 单反相机、PointGrey Grasshopper 50S5C 工业相机（数据由[23]提供）

- 图像劣化的视觉感知差异评价

为了对视觉感知上的颜色失真程度进行定量评价，本论文使用各向异性颜色相似性指

\* 相机灵敏度同色异谱指数（DSCSMI）用于表征某台相机因同色异谱失配现象而产生的颜色失真程度<sup>[78]</sup>。DSCSMI=100 意味着该相机的光谱灵敏度函数与人眼视觉系统的色匹配函数之间完全满足 Luther-Ives 条件<sup>[77]</sup>，此时不存在颜色失真；DSCSMI=50 对应的颜色失真程度大致接近于 D65 光源与暖白荧光灯光源对物体颜色外貌造成的差异<sup>[78]</sup>。

数 (directional statistics-based color similarity index, DSCSI) [211]作为全参考图像质量评价指标对参考图像与同色异谱劣化图像之间的差异程度进行分析。DSCSI-IQA 评价指标在提出时仅用于显示设备的色域映射算法性能评估,但在后续的实验中已经被证实其分数能够较为准确地反映观察者对于两幅待测图像间的视觉感知差异。

将参考图像与测试图像 (即劣化图像) 输入 DSCSI-IQA 模型后, 其输出分数越高, 则表明两幅图像间的视觉感知差异越小; 反之, 则说明测试图像中因颜色失真而导致的劣化越严重。

#### 4.4.2.3 结果与分析

为避免引起不必要的歧义, 本小节将分别对不同测试相机以及不同测试光源所对应的同色异谱现象进行分析。

图 4-37 中展示了使用 Nikon D80 相机在 F8 测试光源下对高光谱图像 “painting” [206] 进行仿真得到的不同种类的劣化图像, 其中图 4-37 (b) ~ (d) 分别表示百分数  $p=0$ 、 $50$  以及  $95$  时对应的物体色同色异谱劣化图像, 即  $\mathcal{I}_{\text{ocm}}^{0\%}$ 、 $\mathcal{I}_{\text{ocm}}^{50\%}$  以及  $\mathcal{I}_{\text{ocm}}^{95\%}$ 。显然, 随着同色异谱程度的增加, 劣化图像与参考图像之间的差异也愈发显著。

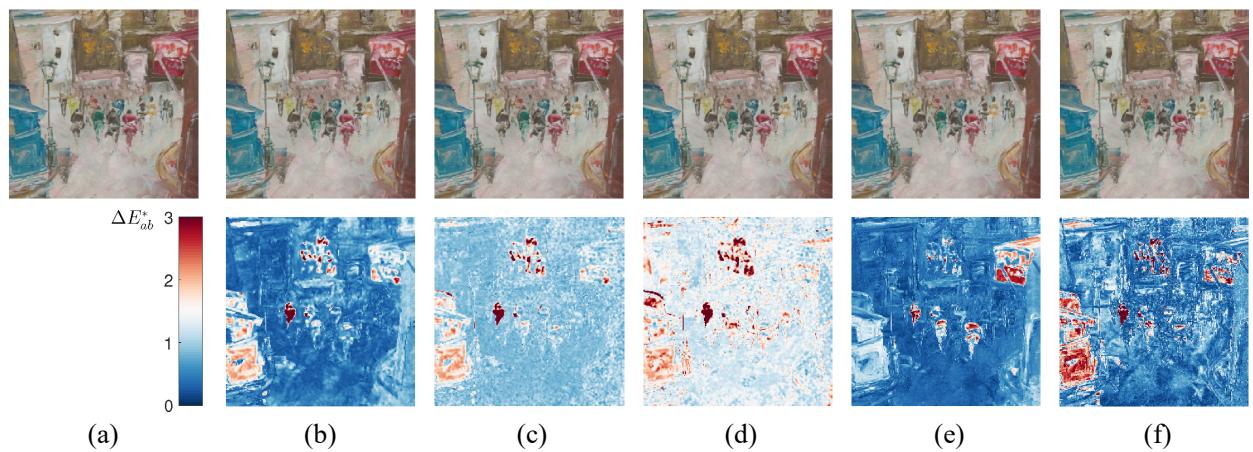


图 4-37 Nikon D80 相机在 F8 测试光源下对高光谱图像 “painting” [206] 仿真得到的劣化图像 (上) 以及对应的色差分布图 (下), 其中 (a) 为参考图像, (b)~(d) 分别对应百分数  $p=0$ 、 $50$  以及  $95$  时的物体色同色异谱劣化图像, (e) 表示 CIECAM02 色貌模型预测图像, (f) 表示光源色同色异谱劣化图像

为使读者对于 DSCSI-IQA 分数与图像劣化程度之间的相关性产生更直观的认识, 除同色异谱劣化图像之外, 我们还额外生成了两类劣化图像并将之与参考图像进行比较。第一类图像表示使用 CIECAM02 色貌模型 [93] 对不同测试光源下拍摄的图像进行对应色

(corresponding color) 预测后得到的图像，即对于某张高光谱图像，首先计算其各个像素在测试光源  $U$  下的 XYZ 三刺激值，接着利用 CIECAM02 色貌模型计算其在参考光源  $C$  下对应色的 XYZ 三刺激值，最后转换至 sRGB 颜色空间中作为仿真图像的驱动值。在调用 CIECAM02 色貌模型时，本论文使用如下参数：

$$\begin{cases} F = 1.0 \\ c = 0.69 \\ N_c = 1.0 \\ L_A = 0.2L_w \end{cases} \quad (4.77)$$

式中  $F$ 、 $c$ 、 $N_c$ 、 $L_A$  分别表示最大适应程度、环境影响非线性参数、色诱导因子以及适应亮度， $L_w$  表示场景中参考白点的亮度。显然，CIECAM02 色貌模型仅接收色度刺激值作为输入，因此仍然不能避免因同色异谱而导致的颜色失真的问题。通过对 CIECAM02 色貌模型预测图像的劣化程度进行分析，我们能够近似地了解人眼视觉系统在进行跨光源颜色感知时存在的颜色失真程度。图 4-37(e) 中展示了 F8 测试光源下使用 CIECAM02 色貌模型进行颜色预测时得到的劣化图像。

第二类图像表示使用错误的标定光源进行颜色校正后得到的图像，即对于在测试光源  $U \in \{\text{CWF, F8, LED-D50, Flash}\}$  下拍摄的图像，我们故意选取一组错误的颜色校正矩阵  $\mathbf{M}(U')$  对其进行校正（详见第 4.4.1.1 小节），其中  $U'$  表示与  $U$  具有相同相关色温的重组日光光源（参见图 4-35）。由于在传统的 ISP 标定流程中往往只会选取若干个具有不同相关色温的典型光源对颜色校正矩阵进行标定，因此在测试阶段对输入图像进行颜色校正时，必然存在一定的概率选取到同色异谱光源所对应的颜色校正矩阵。例如，在标定阶段使用 3000K、5000K、6500K 三种钨丝灯作为标定光源对颜色校正矩阵进行计算，那么在测试阶段，对于一幅在 5000K 荧光灯光源下拍摄的图像，颜色校正模块仍然会调用 5000K 钨丝灯对应的颜色校正矩阵对其进行校正，此时图像中部分物体表面的颜色复现必然将出现错误。通过对这一颜色校正矩阵错误选取的现象进行模拟，我们能够将 ISP 中颜色校正模块可能遇到的光源色同色异谱失配现象与物体色同色异谱失配现象进行简要的比较。图 4-37(f) 中展示了 F8 测试光源下调用 5000K 日光光源对应的颜色校正矩阵时得到的光源色同色异谱劣化图像。

- 不同测试相机的同色异谱颜色失真

图 4-38 中展示了基于 1088 幅测试图像（136 幅高光谱测试图像  $\times$  8 种测试光源）分

析得到的三台测试相机在不同的物体色同色异谱劣化程度下对应的 DSCSI-IQA 分数均值以及标准差。需要注意的是，对于不同的测试相机，其纵坐标分别对应于不同的尺度。作为对比，我们同时也绘制了使用 CIECAM02 色貌模型进行跨光源对应色预测以及使用错误的颜色校正矩阵进行颜色校正时（即光源色同色异谱）对应的 DSCSI-IQA 分数。从图 4-38 中可以观察到，对于两台商用单反相机而言，因物体色同色异谱而引起的颜色失真程度近似等价于使用 CIECAM02 色貌模型对图像进行跨光源颜色复现时产生的失真；当物体色同色异谱的程度不超过中位水平时 ( $p\% \leq 50\%$ )，其对物体颜色复现造成的影响甚至要小于人眼视觉系统进行跨光源颜色感知时所产生的颜色失真。从这一结论来看，当我们使用常见的商用相机对场景中的物体进行跨光源颜色复现时，并不需要过于担心因物体色同色异谱失配现象导致的颜色失真与图像质量劣化。此外，从三台测试相机 DSCSI-IQA 分数的分布趋势上来看，其同色异谱劣化程度与相机灵敏度同色异谱指数 (DSCSMI) 近似成负相关关系，因此在一些不需要定量分析的场合中，我们可以使用 DSCSMI 分数对测试相机的同色异谱程度进行近似估计。

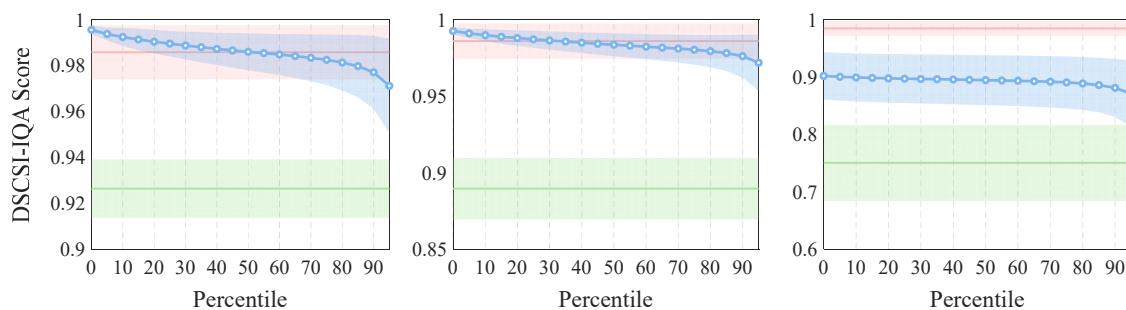


图 4-38 三台测试相机在不同的物体色同色异谱劣化程度下对应的 DSCSI-IQA 分数（蓝色圆圈）。作为对比，红色与绿色实线分别表示使用 CIECAM02 色貌模型进行跨光源对应色预测以及使用错误的颜色校正矩阵进行颜色校正时（即光源色同色异谱）对应的 DSCSI-IQA 分数。浅色阴影面积表示所有测试样本间的 DSCSI-IQA 分数标准差。由左至右：Canon 60D 单反相机、Nikon D80 单反相机、PointGrey Grasshopper 50S5C 工业相机

从图 4-38 中还可以观察到，对于三台测试相机，即使处于物体色同色异谱最为严重的情况下 ( $p\% = 95\%$ )，其产生的劣化程度仍然要小于光源色同色异谱所带来的影响。换言之，在使用颜色校正模型对图像进行跨光源颜色复现时，其（光源相关的）校正参数的选取才是影响图像中颜色复现精度的最主要因素。为缓解这一问题，一方面，在对颜色校正参数进行标定时我们应选用数量更多的标定光源；另一方面，为降低选择错误颜色校正参数的概率，也应尽可能地提高光源颜色估计模型的估计准确度。

需要注意的是，在对物体色同色异谱劣化图像  $\mathcal{I}_{\text{ocm}}^{p\%}$  进行仿真的过程中，我们对其各个像素的 OCMS 进行了近似独立的计算，因此对于  $p\% = 90\%、95\%$  这类同色异谱程度极为严重的情形而言，其仿真结果所反映出的图像劣化程度实际上已经夸大了同色异谱现象在现实世界中所产生的影响——以  $p\% = 95\%$  为例，当且仅当图像中的每一个像素均处于物体色同色异谱最为严重的情况下时才能得到  $\mathcal{I}_{\text{ocm}}^{95\%}$ ，而如此“倒霉”的情况在现实世界中基本不可能出现。换句话说，在现实世界中，图像的劣化程度与同色异谱现象的严重程度之间的变化关系较图 4-38 中绘出的曲线而言，其下降趋势应该更为缓慢一些。

- 不同测试光源的同色异谱颜色失真

在对三台测试相机在不同测试光源下的图像劣化情况进行分析后发现，不同相机 DSCSI-IQA 分数的分布趋势具有高度的一致性，唯一区别仅在于其绝对数值的大小。因此出于简洁起见，这里仅展示了 8 种测试光源下 Nikon D80 单反相机在不同的物体色同色异谱劣化程度下对应的 DSCSI-IQA 分数均值以及标准差，如图 4-39 所示。同样地，作为对比，我们同时也绘制了使用 CIECAM02 色貌模型进行跨光源对应色预测以及使用错误的颜色校正矩阵进行颜色校正时（即光源色同色异谱）对应的 DSCSI-IQA 分数。对于 D50 以及 D100 测试光源，其对应的重组日光光源就是它们自身，因此图 4-39 中这两组测试光源不存在对应的光源色同色异谱分析结果（绿色直线与阴影）；由于重组日光光谱仅适用于相关色温高于 4000K 的情况，因此 A 与 TL84 测试光源同样也不存在对应的光源色同色异谱分析结果。

与图 4-38 中所反映的规律类似，对于不同的测试光源，当  $p\%$  不是太大时，因物体色同色异谱而引起的颜色失真基本与 CIECAM02 色貌模型进行对应色预测时产生的失真程度近似，小于光源色同色异谱引起的颜色失真。同时，从图 4-39 中可以观察到，同色异谱导致的颜色失真程度在不同光源间存在显著的差异。对于 D50 与 D100 测试光源，因物体色同色异谱引起的颜色失真十分轻微，即使在最坏的情况下 ( $p\% = 95\%$ ) 下 DSCSI-IQA 分数均值仍然维持在 0.98 以上。此外，对于 D50 与 D100 测试光源，其样本间的标准差也明显低于其它测试光源，这意味着其颜色失真程度在不同的测试图像之间相对接近。出乎意料的是，具有显色指数  $R_a = 100$  的 A 光源在同色异谱劣化程度上甚至较 F8、LED-D50 等显色指数较低的光源更为严重，由此可以推断，由于物体色同色异谱而引起的颜色失真程度与光源的显色指数之间并不存在明显的相关性。

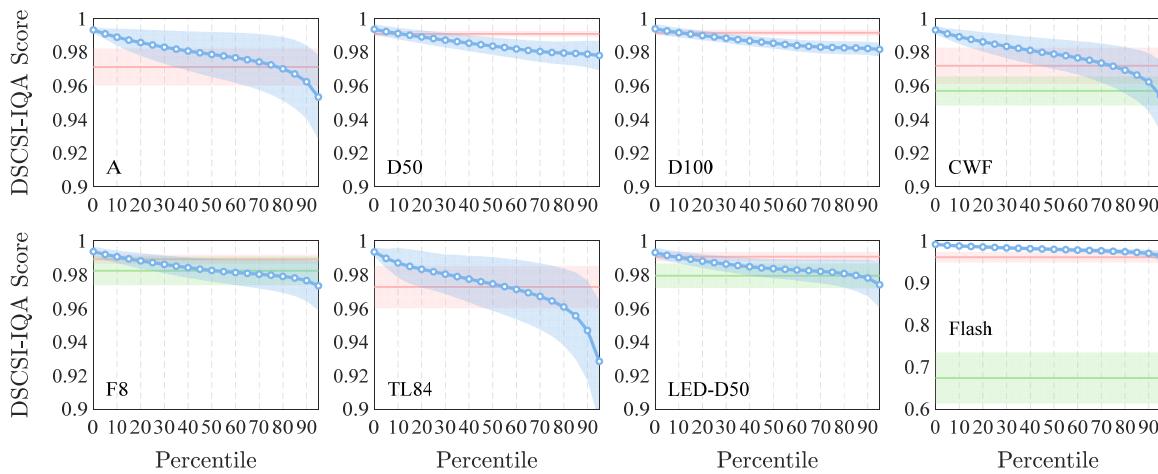


图 4-39 八种测试光源下不同的物体色同色异谱劣化程度对应的 DSCSI-IQA 分数 (蓝色圆圈)。作为对比, 红色与绿色实线分别表示使用 CIECAM02 色貌模型进行跨光源颜色预测以及使用错误的颜色校正矩阵进行颜色校正时 (即光源色同色异谱) 对应的 DSCSI-IQA 分数。浅色阴影面积表示所有测试样本间的 DSCSI-IQA 分数标准差

需要强调的是, 对于 Flash 光源, 其光源色同色异谱对应 DSCSI-IQA 分数 (绿色直线与阴影)要显著地低于其物体色同色异谱以及 CIECAM02 色貌模型颜色预测对应的分数, 同时也显著低于其它 7 种测试光源对应的光源色同色异谱分数, 这是因为 Flash 光源的色品与黑体轨迹偏离较远 ( $D_{uv,Flash} = 0.0101$ , 而通常的“白光”光源所允许的偏离黑体轨迹的最大距离  $D_{uv} = 0.006$ <sup>[190]</sup>), 因此与其同色异谱的重组日光光源实际上具有了略微不同的色品\*, 此时的颜色校正矩阵自然也将原图像 (Flash 光源下拍摄图像) 校正至一个略微“偏色”的颜色外貌。

从图 4-39 中还可以发现, 对于不同的测试光源, 当不存在物体色同色异谱时 ( $p\% = 0\%$ ), 其 DSCSI-IQA 分数基本处于同一水平 ( $\approx 0.996$ )。正如第 4.4.2.1 小节中所述, 当  $p\% = 0\%$  时, 劣化图像  $\mathcal{I}_{ocm}^{0\%}$  与参考图像  $\mathcal{I}^*$  间的误差完全来自于颜色校正模型自身的校正误差, 由此可以推断, 二次项的齐次根式多项式颜色校正模型对于 8 种测试光源的颜色校正性能基本处于同一水平。在这一前提下, 在对不同光源下的颜色失真程度进行分析时, 因颜色校

\* “同色异谱”中的“色”指的是色品而非相关色温, 因此严格地说, 此时的情形并不能算同色异谱。然而, 在颜色校正模块实际工作时, 目标光源与标定光源相关色温相同而色品不同, 甚至是相关色温不同的情形也时有出现。由于本小节中我们的研究重点在于对实际 ISP 流程中可能出现的颜色失真进行分析, 因此允许适当放宽关于“同色异谱”的定义。

正模型精度不足而造成的残留误差在一定程度上可以相互抵消，因此图 4-39 中所反映的规律并不会随着颜色校正模型的改变而发生变化。

## 4.5 图像后处理

为了使最终的输出图像具有与人眼视觉系统感知更加贴近的颜色外貌，本节使用 CIECAM02 色貌模型中的 CAT02 色适应模型在图像后处理模块中对图像的整体颜色外貌进行调节<sup>[97]</sup>。在完成该颜色外貌调节后，还将使用双边滤波算法<sup>[212]</sup>对输出图像进行简单的降噪操作，并利用 sRGB 非线性函数<sup>[169]</sup>作为阶调映射模型将线性图像转换至最终可用于显示的非线性图像，至此完成本论文中全部的图像信号处理步骤。

### 4.5.1 基于色适应变换模型的图像颜色外貌调节算法

色适应是颜色恒常性的具体表现形式之一。色貌模型中通常使用符号  $D$  对人眼视觉系统的色适应程度进行表示： $D=1$  表示完全色适应，即实现了理想的颜色恒常性； $D=0$  表示无任何适应能力，此时人眼视觉系统等价于一台测色仪器，即只对物体表面的光度、色度信息进行记录，而未进行任何主观的调节。在白平衡校正与颜色校正章节中，我们将数码相机 ISP 流程中的颜色复现过程定义为“将任意光源下拍摄的图像完全校正至参考光源下的颜色外貌”，即默认  $D=1$ 。然而，人眼视觉系统中的颜色恒常性并不总是完美的，在某些色品严重偏离中性点的光源的照明下，或在某些低照度的观察环境中，人眼并不能完全无视照明体的存在而判断出物体真实的颜色。此时，若仍将  $D=1$  作为数码相机的颜色复现目标，即对任何照明光源都无差别地使用统一的参考光源作为颜色复现的目标，则有可能对最终的 ISP 输出图像产生如下弊端：1) 输出图像过“白”，场景中物体颜色的还原不符合观察者的真实视觉感知，对于身处拍摄场景中的拍摄者尤甚；2) 光源对场景氛围的渲染作用完全被抑制；3) 将偏色较严重的光源强行校正至参考光源，会增加空间非均匀性校正、颜色校正等模块的处理难度，甚至导致最终的图像质量出现劣化。为改善这一问题，我们利用色适应模块计算出一组色适应变换后增益系数（chromatic adaptation transform post-gains）并由此对图像进行整体颜色外貌调节。

#### 4.5.1.1 色适应后增益系数的计算

为了避免对 ISP 流程中原有模块的工作流程产生影响，我们将色适应模块置于空间非均匀性校正、白平衡校正以及颜色校正模块之后。

色适应模块以非均匀性校正后的线性图像作为输入，通过对其中响应值进行分析从而估计出拍摄场景中的白场亮度  $L_w$ （以  $\text{cd}/\text{m}^2$  为单位），并将其数值的 20% 作为该场景中的适应场亮度  $L_A$ <sup>[93]</sup>；同时，利用自动白平衡校正模块中估计出的通道增益系数  $\mathcal{G}_r$ 、 $\mathcal{G}_b$ ，可对拍摄场景中的光源（即待适应光源）三刺激值  $[X_w, Y_w, Z_w]$  进行估计。在得到  $L_A$  与  $[X_w, Y_w, Z_w]$  后，色适应模块对场景中的白场在参考光源下的对应色进行预测，并将其三刺激值与参考光源的三刺激值进行比较，从而得到最终的调制系数。由于色适应模块对图像颜色外貌的调节过程与白平衡校正模块中使用一组增益系数对图像三通道进行独立调节的过程相类似，因此本论文将色适应模块输出的调制系数称为色适应变换后增益系数，并使用符号  $\mathcal{G}_r^{\text{CAT}}$ 、 $\mathcal{G}_g^{\text{CAT}}$ 、 $\mathcal{G}_b^{\text{CAT}}$  进行表示。与白平衡校正中的操作类似，本节将 **G** 通道对应的色适应后增益系数固定为 1，仅使用  $\mathcal{G}_r^{\text{CAT}}$  与  $\mathcal{G}_b^{\text{CAT}}$  对图像进行全局颜色外貌调节：

$$\begin{cases} \mathcal{G}_r^{\text{CAT}} \leftarrow \mathcal{G}_r^{\text{CAT}} / \mathcal{G}_g^{\text{CAT}} \\ \mathcal{G}_g^{\text{CAT}} = 1 \\ \mathcal{G}_b^{\text{CAT}} \leftarrow \mathcal{G}_b^{\text{CAT}} / \mathcal{G}_g^{\text{CAT}} \end{cases} \quad (4.78)$$

图 4-40 中展示了色适应模块在 ISP 流程中与其它模块的交互逻辑关系。

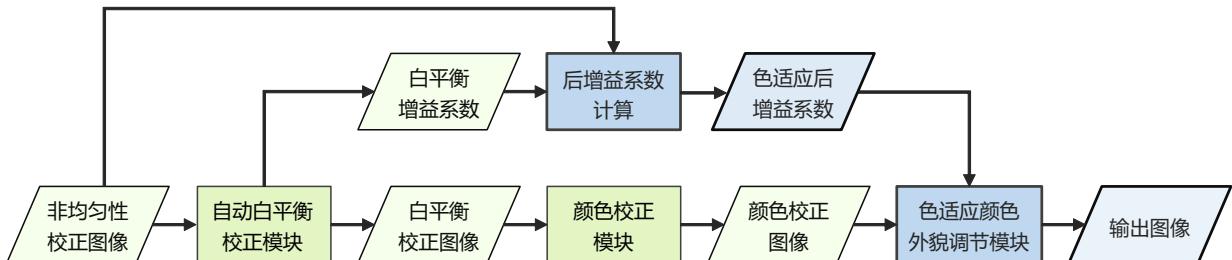


图 4-40 基于色适应变换的图像颜色外貌调节模块（蓝色）与 ISP 流程中其它模块（绿色）的交互逻辑关系

具体来说，色适应模块采用如下步骤对颜色校正后的图像颜色外貌进行调节：

- 1) 利用原始响应值预测模型对拍摄场景中白场亮度  $L_w$  进行估计，并将该亮度值的 20% 作为 CAT02 色适应变换模型中的适应场亮度。关于拍摄场景中白场亮度估计的具体原理及实现过程请参阅第 5.3.1 小节。
- 2) 利用自动白平衡校正模块得到的白平衡增益系数  $\mathcal{G}_r$ 、 $\mathcal{G}_b$  计算经白平衡校正后的光源 RGB 响应值：

$$\begin{cases} D_{\text{wb},r}^U = \mathcal{G}_r^C / \mathcal{G}_r^U \\ D_{\text{wb},g}^U = \mathcal{G}_g^C / \mathcal{G}_g^U = 1 \\ D_{\text{wb},b}^U = \mathcal{G}_b^C / \mathcal{G}_b^U \end{cases} \quad (4.79)$$

3) 将  $D_{\text{wb},r}^U$ 、 $D_{\text{wb},g}^U$ 、 $D_{\text{wb},b}^U$  近似地视为线性 sRGB 响应值，并使用经验公式 (4.63) 对拍摄场景中光源的 XYZ 三刺激值进行估算：

$$\begin{bmatrix} \hat{X}^U \\ \hat{Y}^U \\ \hat{Z}^U \end{bmatrix} = \begin{bmatrix} 0.413 & 0.213 & 0.019 \\ 0.358 & 0.715 & 0.119 \\ 0.180 & 0.072 & 0.950 \end{bmatrix} \begin{bmatrix} D_{\text{wb},r}^U \\ D_{\text{wb},g}^U \\ D_{\text{wb},b}^U \end{bmatrix} \quad (4.80)$$

4) 利用 CAT02 色适应变换模型 (2.31) 计算参考光源下  $[\hat{X}^U, \hat{Y}^U, \hat{Z}^U]$  的对应色  $\mathbf{X}_c$ ：

$$\mathbf{X}_c = [\hat{X}_c^U, \hat{Y}_c^U, \hat{Z}_c^U] = f_{\text{CAT02}}(\mathbf{X}, \mathbf{X}_w, \mathbf{X}_w^{\text{ref}}, L_A, B) \quad (4.81)$$

式中  $\mathbf{X} = \mathbf{X}_w = [\hat{X}^U, \hat{Y}^U, \hat{Z}^U]$ ,  $\mathbf{X}_w^{\text{ref}} = [0.9505, 1, 1.0888]$  表示参考光源 D65 的 XYZ 三刺激值，背景相对亮度等级  $B$  采用平均亮度 (average)。

5) 计算色适应后光源的线性 sRGB 值，并将其与参考光源的线性 sRGB 值之比作为色适应后增益系数：

$$\begin{cases} \mathcal{G}_r^{\text{CAT}} = \hat{R}_c^U / R^C \\ \mathcal{G}_g^{\text{CAT}} = \hat{G}_c^U / G^C \\ \mathcal{G}_b^{\text{CAT}} = \hat{B}_c^U / B^C \end{cases} \quad (4.82)$$

其中

$$\begin{bmatrix} \hat{R}_c^U \\ \hat{G}_c^U \\ \hat{B}_c^U \end{bmatrix} = \begin{bmatrix} 0.413 & 0.213 & 0.019 \\ 0.358 & 0.715 & 0.119 \\ 0.180 & 0.072 & 0.950 \end{bmatrix}^{-1} \begin{bmatrix} \hat{X}_c^U \\ \hat{Y}_c^U \\ \hat{Z}_c^U \end{bmatrix}, \quad \begin{bmatrix} R^C \\ G^C \\ B^C \end{bmatrix} = \begin{bmatrix} 0.413 & 0.213 & 0.019 \\ 0.358 & 0.715 & 0.119 \\ 0.180 & 0.072 & 0.950 \end{bmatrix}^{-1} \begin{bmatrix} X^C \\ Y^C \\ Z^C \end{bmatrix} \quad (4.83)$$

6) 使用式 (4.78) 对色适应后增益系数进行归一化。

7) 将颜色校正后的图像各通道乘上对应的色适应后增益系数，完成颜色外貌调节。

图 4-41 与图 4-42 展示了色适应模块对于测试图像颜色外貌的影响。当拍摄场景中环境亮度较低且照明光源的色品明显偏离参考光源时，色适应模块能够自适应地对图像的整体颜色外貌进行调节，从而使得输出图像与观察者的视觉感知之间具有更好的一致性；随着环境亮度的提高，人眼视觉系统的色适应程度  $D$  逐渐增大，因此色适应模块对图像颜色外貌调节的力度逐渐下降，直至  $D=1$  时不再对输入图像进行任何调节。

需要说明的是，对于一些存在饱和像素的输入图像，若  $\mathcal{G}_r^{\text{CAT}}$  与  $\mathcal{G}_b^{\text{CAT}}$  二者之中存在小于

1的值，则有可能导致色适应调节后图像中的高光区域出现偏色。为避免引入这一不必要的失真，一种更稳妥的做法是，在自动白平衡校正模块完成光源颜色估计任务后，将一份未经白平衡校正的图像拷贝至数据缓存区之中，同时继续执行后续的ISP流程。待色适应模块完成后增益系数的计算后，将后增益系数与白平衡校正增益系数进行合并，以此作为“色适应调制后的增益系数”对缓存区中的图像进行白平衡校正，即

$$\begin{cases} \mathcal{G}_r \leftarrow \mathcal{G}_r \times \mathcal{G}_r^{\text{CAT}} \\ \mathcal{G}_b \leftarrow \mathcal{G}_b \times \mathcal{G}_b^{\text{CAT}} \end{cases} \quad (1.84)$$

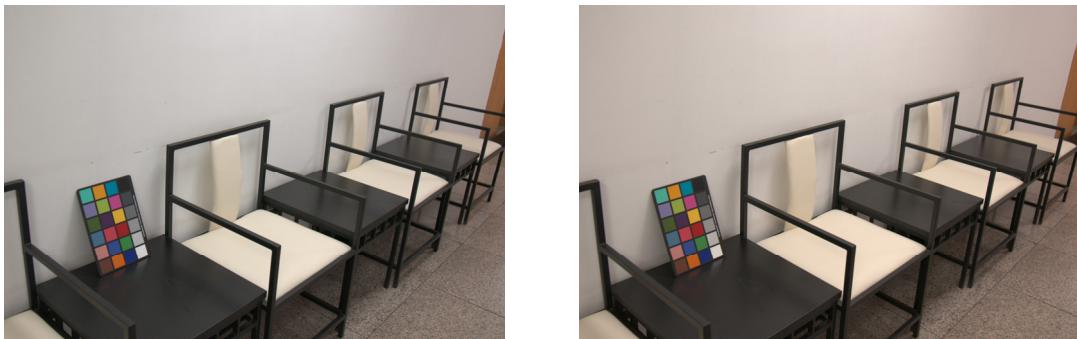


图 4-41 完全校正至参考光源的图像（左）与色适应模块调节颜色外貌后的图像（右），为便于可视化起见两张图像均已经过 gamma 校正。该场景中光源的真实相关色温约为 4400K，白场亮度约  $75\text{cd}/\text{m}^2$ （对应于图像上侧的白墙），观察者身处该场景中时能够明显感觉到周围物体呈现出暖色调



图 4-42 完全校正至参考光源的图像（左）与色适应模块调节颜色外貌后的图像（右），为便于可视化起见两张图像均已经过 gamma 校正。该场景中光源的真实相关色温约为 6300K，白场亮度约  $1200\text{cd}/\text{m}^2$ 。随着环境光亮度的提高，色适应模块对图像颜色外貌的所起的调节作用逐渐降低

## 4.6 本章小结

为了获得与人眼视觉感知具有一致颜色外貌的输出图像，数码相机需要通过 ISP 流程

对传感器采集的原始图像进行一系列的分析与处理。本章依据 ISP 流程中图像信号流传输顺序，分别对空间非均匀性校正、白平衡校正、颜色校正以及图像后处理等与颜色复现紧密相关的图像处理模块进行了介绍。

在空间非均匀性校正中，提出了基于 B-样条曲面的数字图像空间非均匀性校正算法，该算法通过对光源估计模块中得到的光源颜色信息进行分析，自适应地选取合适的校正参数对输入图像进行空间非均匀性校正。

在颜色恒常性与自动白平衡校正中，分别介绍了一种基于先验假设的光源颜色估计算法 GMap-NPStat 以及基于一种基于学习的光源颜色估计算法 ReWNet，其中 GMap-NPStat 算法简单可靠，且在模型性能与复杂度之间取得了良好的平衡，十分适合部署于现阶段移动终端内的数码相机之中；ReWNet 算法基于深度卷积网络进行设计，在性能上与目前最佳的颜色恒常性算法具备一定的竞争力，且相比于现有的基于卷积神经网络的模型而言具有很强的可解释性。

在颜色校正中，基于根式多项式模型对现有的颜色校正参数标定方法进行了改进，在不影响执行效率的前提下提升了数码相机的跨光源颜色复现精度；同时，基于大量的真实数据，对数码相机颜色复现过程中的物体色同色异谱失配现象进行了定量分析，最终得到结论：物体色同色异谱并不是造成数字成像设备颜色失真的主要原因，相比于物体色同色异谱，我们更应该关注相机光谱灵敏度函数与标准观察者色匹配函数之间的近似关系，以及如何尽可能避免光源色同色异谱对颜色复现带来的影响。

在图像后处理中，提出了一种基于 CAT02 色适应变换模型的图像颜色外貌调节算法，该算法能够根据场景的亮度信息以及光源颜色信息自适应地对图像的整体颜色外貌进行调节，从而使得输出图像更加贴合观察者在拍摄场景中实际感知到的颜色外貌。

## 第五章 原始响应值预测模型应用实例

基于本论文第四章中关于 ISP 流程中各模块实现方法与原理的阐述，本章将以 Nikon D3x 相机作为范例，对数码相机原始响应值预测模型在自动白平衡校正、颜色校正、同色异谱失配分析以及图像后处理等模块的标定任务中所起的作用进行展示，其流程结构与交互关系如图 5-1 所示。

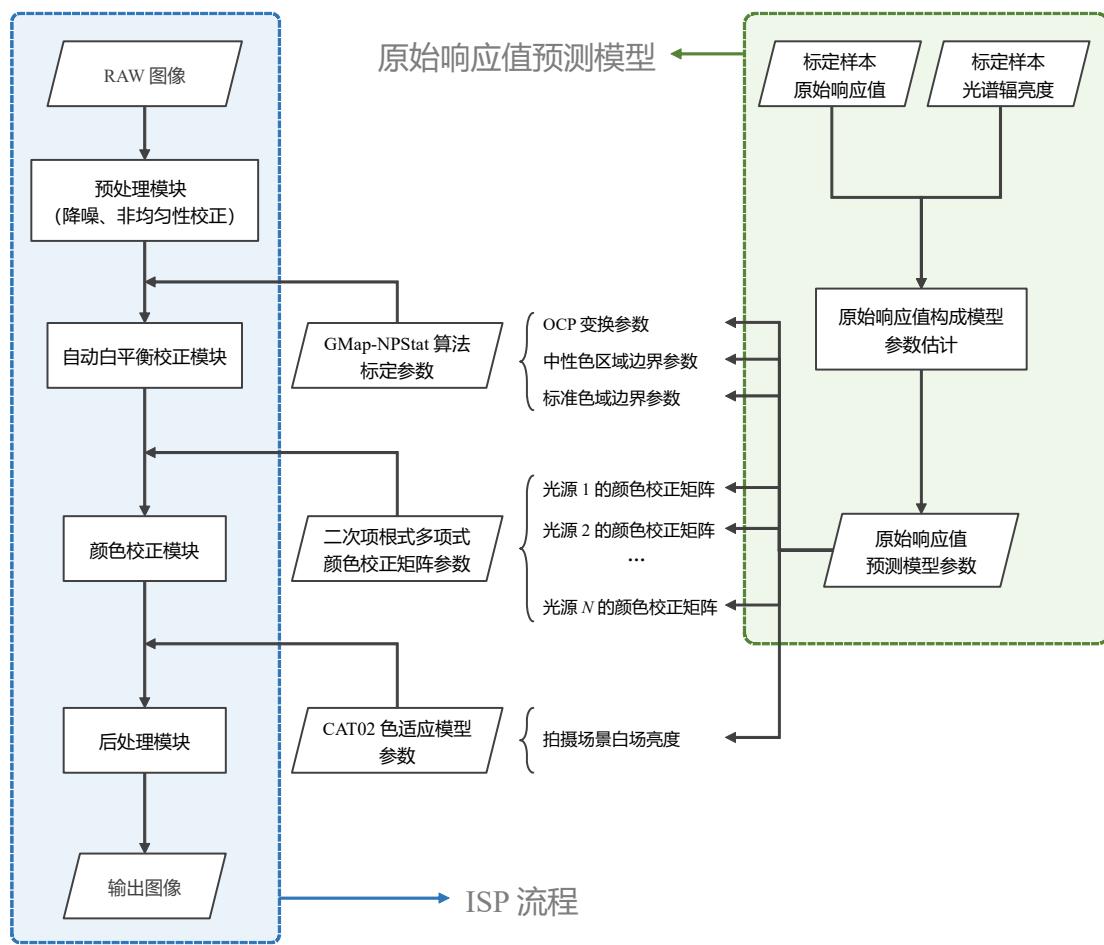


图 5-1 原始响应值预测模型在 ISP 流程中的应用实例

### 5.1 原始响应值预测模型在自动白平衡校正模块参数标定中的应用

#### 5.1.1 相机 RGB 颜色空间至正交色度平面的变换

在第 4.3.1.1 小节中，我们要求正交色度平面能够满足以下两点性质：

- 1) 当光源的相关色温发生变化时，该平面内的色度坐标仅沿水平方向发生变化；

2) 当光源的相关色温不变而色品坐标在等色温线上变化时, 该平面内的色度坐标仅沿竖直方向发生变化。

借助原始响应值预测模型, 我们能够十分方便地对相机 RGB 颜色空间至正交色度平面坐标变换关系中的相关参数进行确定。

首先, 参考[213], 使用如下公式去除响应值中的亮度信息:

$$\begin{cases} X = \log \frac{w_r D_r}{w_r D_r + w_g D_g + w_b D_b} \\ Y = \log \frac{w_b D_b}{w_r D_r + w_g D_g + w_b D_b} \end{cases} \quad (5.1)$$

式中  $w_k, k \in \{r, g, b\}$  表示三通道的权重系数, 且有  $w_r + w_g + w_b = 1$ 。

在确定合适的权重系数之前, 不妨先令  $w_r = w_g = w_b = \frac{1}{3}$ , 此时利用原始响应值预测模型, 能够模拟出 3000K~12000K 黑体在该  $(X, Y)$  平面上的坐标分布, 如图 5-2 所示。通过对图 5-2 中的轨迹进行直线拟合, 能够得到该轨迹相对于  $X$  轴的偏转角, 记为  $\theta_0$ 。

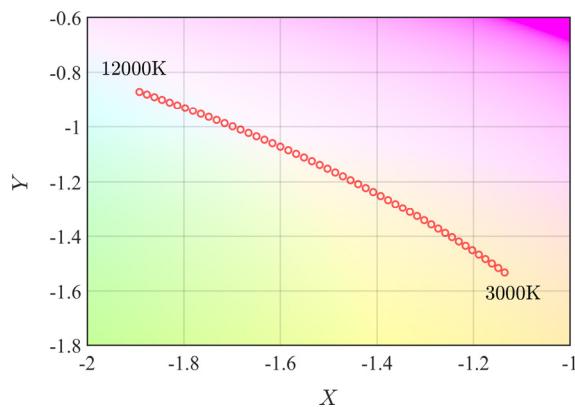


图 5-2 3000K 至 12000K 黑体在 Nikon D3x  $(X, Y)$  平面上的坐标分布 (本例中  $w_r = w_g = w_b = \frac{1}{3}$ ), 其中共包含 60 个均匀分布于色温倒数轴上的样本点

为满足性质 1, 首先对  $XY$  平面进行平移, 使得 D65 光源对应的坐标成为  $(X, Y)$  平面内的原点  $[X_0, Y_0]$ , 同时使用一个旋转矩阵

$$\mathbf{R} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (5.2)$$

对  $(X, Y)$  平面上的坐标进行旋转变换, 使得在旋转后的坐标平面 (记为  $(X_{rot}, Y_{rot})$  平面) 内黑体坐标尽可能沿水平方向分布。这一条件可通过最小化各样本在竖直方向上的方差来

实现：

$$\left[ w_r^*, w_g^*, w_b^*, \theta^* \right] = \arg \min \left\{ \text{var} \left[ \mathbf{Y}_{rot}(w_r, w_g, w_b, \theta) \right] \right\} \quad (5.3)$$

式中  $\text{var}(\cdot)$  表示计算样本方差， $\mathbf{Y}_{rot}$  表示由各样本点旋转后的  $Y_{rot}$  坐标构成的向量，有

$$\begin{bmatrix} X_{rot}(w_r, w_g, w_b, \theta) \\ Y_{rot}(w_r, w_g, w_b, \theta) \end{bmatrix} = \mathbf{R}(\theta) \cdot \begin{bmatrix} X(w_r, w_g, w_b) - X_0(w_r, w_g, w_b) \\ Y(w_r, w_g, w_b) - Y_0(w_r, w_g, w_b) \end{bmatrix} \quad (5.4)$$

我们使用  $\theta = \theta_0$  以及  $w_r = w_g = w_b = \frac{1}{3}$  作为初始值对式 (5.3) 进行最小值搜索，对于 Nikon D3x 相机，最终得到的各个最优参数分别为  $w_r^* = 0.2533$ ,  $w_g^* = 0.4935$ ,  $w_b^* = 0.2532$ ,  $X_0^* = -2.0285$ ,  $Y_0^* = -1.3957$  以及  $\theta^* = 40.85^\circ$ 。图 5-3 展示了经平移与旋转后的黑体坐标在  $(X_{rot}, Y_{rot})$  平面上的分布情况。

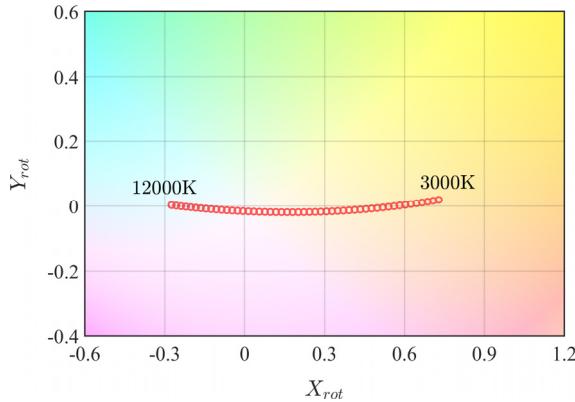


图 5-3 3000K 至 12000K 黑体在 Nikon D3x  $(X_{rot}, Y_{rot})$  平面上的坐标分布

为满足性质 2，首先利用光谱可调的多通道 LED 光源<sup>[122]</sup>生成一组相关色温均为 6500K 但色品在  $-0.02 \leq Duv \leq 0.02$ <sup>[161]</sup> 等色温线上均匀分布的光谱功率分布函数。利用式(5.1)、(5.4)，可得到这组光源在  $(X_{rot}, Y_{rot})$  平面内的坐标分布，如图 5-4 所示。

我们使用一个“修剪 (shearing)”矩阵

$$\mathbf{S} = \begin{bmatrix} 1 & \sigma \\ 0 & 1 \end{bmatrix} \quad (5.5)$$

将  $(X_{rot}, Y_{rot})$  平面坐标变换至正交色度平面  $(X_{orth}, Y_{orth})$  内，使得具有不同  $Duv$  的 LED 光源坐标在  $(X_{orth}, Y_{orth})$  平面内尽可能沿竖直方向分布。这一条件可通过最小化各样本在水平方向上的方差来实现：

$$\sigma^* = \arg \min \left\{ \text{var} \left[ \mathbf{X}_{orth}(\sigma) \right] \right\} \quad (5.6)$$

式中  $\mathbf{X}_{orth}$  表示由各 LED 光源经“修剪”后的  $X_{orth}$  坐标构成的向量，有

$$\begin{bmatrix} X_{orth}(\sigma) \\ Y_{orth}(\sigma) \end{bmatrix} = \mathbf{S}(\sigma) \cdot \begin{bmatrix} X_{rot} \\ Y_{rot} \end{bmatrix} \quad (5.7)$$

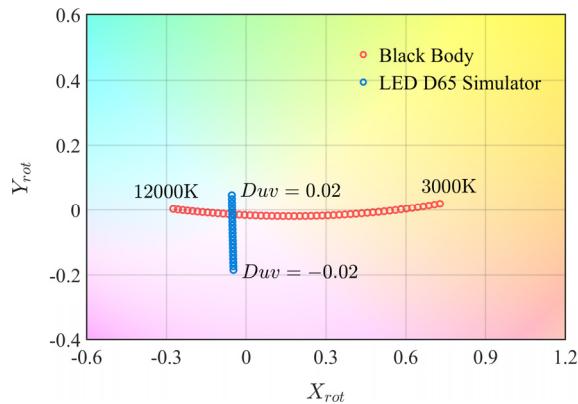


图 5-4 6500K 相关色温、 $-0.02 \leq Duv \leq 0.02$  的一组多通道 LED 光源在 Nikon D3x  $(X_{rot}, Y_{rot})$  平面上的坐标分布（蓝色样本点）

对于 Nikon D3x 相机，最终得到的最优修剪系数  $\sigma^* = -0.0453$ 。图 5-5 展示了经“修剪”后的 LED 光源坐标在  $(X_{orth}, Y_{orth})$  平面上的分布情况。

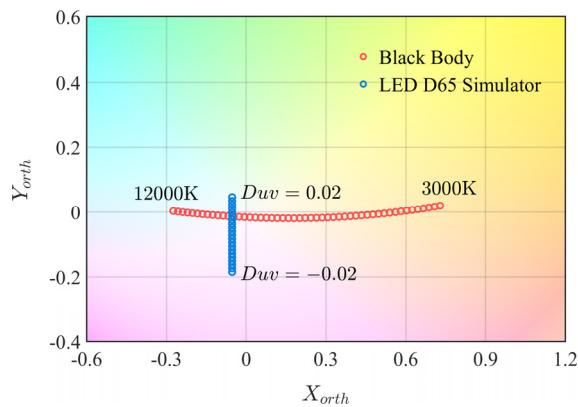


图 5-5 多通道 LED 光源（蓝色样本点）在 Nikon D3x 正交色度平面上的坐标分布

### 5.1.2 NPStat 算法中中性色区域的预设定

由式 (4.23) 可知，不同的中性色区域掩模  $\mathbf{M}_n$  将产生不同的候选中性色像素直方图，因此关于中性色区域范围的选取将直接影响到最终光源颜色估计的准确度。理想的中性色区域既应该具有一定的泛化能力，能够覆盖常见拍摄场景中不同种类的光源，同时也应该具有一定的容错性与抗干扰能力，能够有效地排除伪中性色（例如淡蓝、淡黄色物体表面）以及图像噪声对于光源估计结果的干扰。

本论文中，为了确定合适的中性色区域范围，首先利用原始响应值预测模型对 248 种常见光源<sup>[196, 210]</sup>对应的原始 RGB 响应值进行了仿真，其中包括 50 种色温 3000K-12000K 的黑体辐射源、51 种色温 4000K-12000K 重组日光、6 种 CIE 推荐标准照明体、26 种荧光光源、5 种高压气体放电灯以及 110 种白光 LED 光源，其色品坐标在 CIE1931 xy 色品平面以及正交色度平面上的分布如图 5-6 所示。

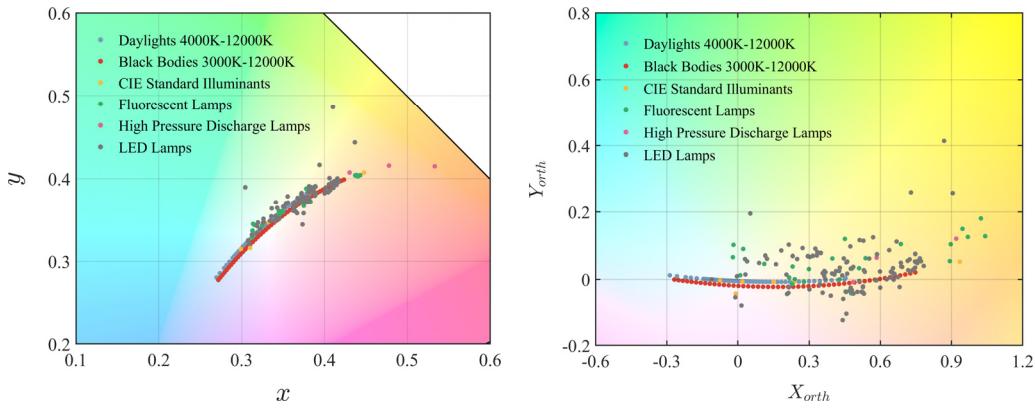


图 5-6 248 种常见光源的色度坐标在 CIE1931 xy 色品平面（左）以及 Nikon D3x 正交色度平面（右）上的分布情况

在手动排除色温过低或色品坐标距离原点过远的部分光源后，参考[213]，我们使用正交色度平面内一个六边形对中性色区域进行表示，如图 5-7 所示。在实际应用中，也可针对不同种类的光源预先设定多个中性色区域，从而缩小每个中性色区域的覆盖范围，提高光源估计算法的容错性。例如，若已知当前图像中的光源属于日光类型（可由场景分类算法预先判断），则中性色区域无需覆盖位于  $Y_{orth} > 0.02$  区域内的荧光光源或 LED 光源，从而大幅降低中性色统计算法对于浅绿色像素的误统计概率（白光下的浅绿色物体与浅绿色光源下的中性色物体对应的响应值均将落入该区域内）。

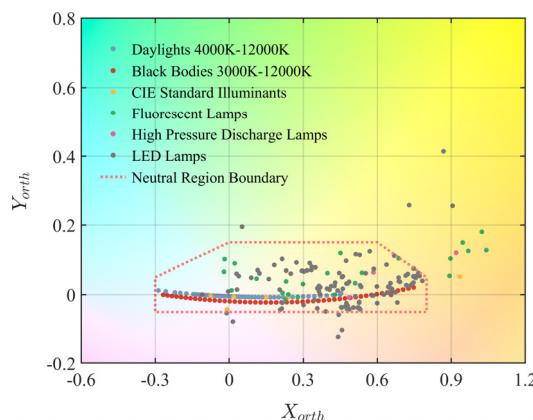


图 5-7 Nikon D3x 正交色度平面内手动选取的中性色区域（红色虚线六边形）

### 5.1.3 二维色域映射算法中标准色域的计算

正如前文所述，为确定某一台相机的标准色域，需要获得该相机对于所有可能存在的物体表面在参考光源下的响应值坐标。显然，在参考光源下对具有不同光谱反射比的物体表面样本进行拍摄并获取其响应值并不是一种可行的方案，一方面，收集并妥善保存数以万计的颜色样本将耗费大量的人力与物力，另一方面，对于每一批甚至每一台测试相机都进行标准色域的标定需要耗费大量的时间，无法应对批量化的生产需求。

为了通过仿真的方法对测试相机的标准色域进行确定，本论文使用 9270 组实测的物体表面光谱反射比函数作为数据源，并利用相机原始响应值预测模型对其在参考光源下的响应值进行计算。9270 组光谱反射比数据共包括 1269 种无光泽的 Munsell 色卡<sup>[214]</sup>、1600 种带光泽的 Munsell 色卡<sup>[214]</sup>、289 种 IT8.7 标准色样<sup>[215]</sup>、1328 种植物表面<sup>[216]</sup>、803 种印刷品与纸制品<sup>[217]</sup>，808 种标准色卡表面、1995 种 Simon Fraser University 实测数据<sup>[218]</sup>以及 1178 种 University of Eastern Finland 实测数据<sup>[219]</sup>。图 5-8 (a) 绘制了其中 200 种具有高饱和颜色的物体表面的光谱反射比函数。

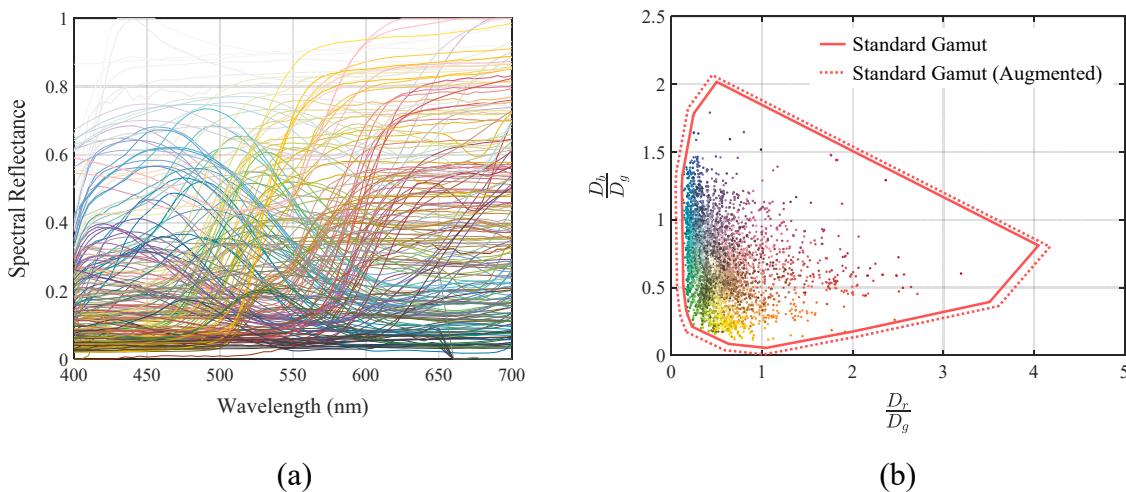


图 5-8 (a) 部分用于生成标准色域的高饱和物体表面的实测光谱反射比函数；(b) 高饱和物体表面样本在 D65 光源下的正交色度平面坐标分布

使用 D65 光源作为参考光源，利用原始响应值预测模型，可获得这 9270 种物体表面对应的 RGB 原始响应值。将这些响应值转换至  $rb$  平面并计算包围这些散点的最小凸集，即可获得测试相机在  $rb$  平面上的标准色域范围，如图 5-8 (b) 中红色实线所示。最后，考虑到所收集的 9270 种光谱反射比函数相对于实际拍摄场景中所有可能出现的物体表面而言难免有所遗漏，我们人为地将该最小凸集向外扩充了 5% 并以此作为最终的标准色域范

围，如图 5-8 (b) 中红色虚线所示。

## 5.2 原始响应值预测模型在颜色校正模块参数标定中的应用

### 5.2.1 二次项根式多项式颜色校正矩阵的标定

对于光谱反射比函数为  $\rho(\lambda)$  的目标样本，联立式 (2.9) 与 (2.12)，可求得其在参考光源  $C$  下的色度三刺激值

$$\mathbf{X}^C = \mathbf{\rho} \cdot \text{diag}(\mathbf{P}^C) \cdot \mathbf{S}_{\text{cmf}}^T \quad (5.8)$$

式中  $\mathbf{X}^C = [X^C, Y^C, Z^C]$  表示三刺激值向量； $\mathbf{\rho} = [\rho(380), \dots, \rho(780)]$  表示光谱反射比函数对应的  $1 \times N$  行向量 ( $N$  为波长采样数)； $\mathbf{P}^C = [P^C(380), \dots, P^C(780)]$  表示参考光源光谱功率分布函数对应的  $1 \times N$  行向量； $\text{diag}(\cdot)$  表示将向量转换为对角矩阵； $\mathbf{S}_{\text{cmf}}^T$  表示  $N \times 3$  的标准观察者色匹配函数矩阵，每一列对应  $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 、 $\bar{z}(\lambda)$  中的一组色匹配函数（参见式 (2.15)）。

同理，联立式 (2.9) 与 (3.81)，还可得到该目标样本在待标定光源  $U$  下的相机原始响应值：

$$\mathbf{D}^U = Tg\kappa(\Delta\lambda)\mathbf{\rho} \cdot \text{diag}(\mathbf{P}^U) \cdot \mathbf{S}^T \quad (5.9)$$

式中  $\mathbf{D}^U = [D_r^U, D_g^U, D_b^U]$  表示该样本的原始响应值向量； $\mathbf{P}^U$  表示未知光源光谱功率分布函数对应的  $1 \times N$  行向量； $\mathbf{S}^T$  表示  $N \times 3$  的相机光谱灵敏度函数矩阵，每一列对应一个颜色通道的相机光谱灵敏度函数（参见式 (3.70)）。

我们使用光源  $U$  对应的白平衡增益系数  $\mathcal{G}_r^U$ 、 $\mathcal{G}_b^U$  对原始响应值  $\mathbf{D}^U$  进行白平衡校正（ $\mathcal{G}_g^U$  已默认为 1），并将白平衡校正后的响应值向量记为  $\mathbf{D}_{\text{wb}}^U$ ，即

$$\begin{aligned} \mathbf{D}_{\text{wb}}^U &= \left[ D_{\text{wb},r}^U, D_{\text{wb},g}^U, D_{\text{wb},b}^U \right] \\ &= \left[ \mathcal{G}_r^U D_r^U, D_g^U, \mathcal{G}_b^U D_b^U \right] \end{aligned} \quad (5.10)$$

对于待标定光源  $U$ ，二次项齐次根式多项式模型使用一个  $6 \times 3$  颜色校正矩阵（color correction matrix） $\mathbf{M}(U)$  将白平衡校正后的响应值  $\mathbf{D}_{\text{wb}}^U$  变换至参考光源  $C$  下的色度三刺激值  $\mathbf{X}^C$ 。该颜色校正矩阵的标定过程如下：

- 1) 对于光谱反射比函数分别为  $\rho_1(\lambda), \dots, \rho_M(\lambda)$  的  $M$  个训练样本，根据式 (5.8) 计算其在参考光源  $C$  下的色度三刺激值向量，并将  $M$  组向量进行逐行排列，构成  $M \times 3$  的三刺激值矩

阵  $\mathbf{X}'^C$  :

$$\mathbf{X}'^C = \begin{bmatrix} \mathbf{X}_1^C \\ \vdots \\ \mathbf{X}_M^C \end{bmatrix} = \begin{bmatrix} X_1^C & Y_1^C & Z_1^C \\ \vdots & \vdots & \vdots \\ X_M^C & Y_M^C & Z_M^C \end{bmatrix} \quad (5.11)$$

2) 对于  $M$  个训练样本, 根据式 (5.9) 与 (5.10) 计算其白平衡校正后的响应值向量  $\mathbf{D}_{\text{wb}}^U$ , 并对其进行二次项根式多项式扩展, 得到  $1 \times 6$  的根式多项式向量  $\tilde{\mathbf{D}}_{\text{wb}}^U$ :

$$\tilde{\mathbf{D}}_{\text{wb}}^U = \left[ D_{\text{wb},r}^U, D_{\text{wb},g}^U, D_{\text{wb},b}^U, \sqrt{D_{\text{wb},r}^U D_{\text{wb},g}^U}, \sqrt{D_{\text{wb},r}^U D_{\text{wb},b}^U}, \sqrt{D_{\text{wb},g}^U D_{\text{wb},b}^U} \right] \quad (5.12)$$

其中曝光时间  $T$  与综合增益系数  $\mathbf{g}$  应根据光源  $U$  的发光强度进行合适的调节, 以确保白平衡校正后的  $M$  组响应值均不出现饱和且具有足够高的信噪比。完成扩展后, 将  $M$  组  $1 \times 6$  向量进行逐行排列, 构成  $M \times 6$  的根式多项式响应值矩阵  $\tilde{\mathbf{D}}'_{\text{wb}}^U$ :

$$\begin{aligned} \tilde{\mathbf{D}}'_{\text{wb}}^U &= \begin{bmatrix} \tilde{\mathbf{D}}_{\text{wb},1}^U \\ \vdots \\ \tilde{\mathbf{D}}_{\text{wb},M}^U \end{bmatrix} \\ &= \begin{bmatrix} D_{\text{wb},1,r}^U & D_{\text{wb},1,g}^U & D_{\text{wb},1,b}^U & \sqrt{D_{\text{wb},1,r}^U D_{\text{wb},1,g}^U} & \sqrt{D_{\text{wb},1,r}^U D_{\text{wb},1,b}^U} & \sqrt{D_{\text{wb},1,g}^U D_{\text{wb},1,b}^U} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ D_{\text{wb},M,r}^U & D_{\text{wb},M,g}^U & D_{\text{wb},M,b}^U & \sqrt{D_{\text{wb},M,r}^U D_{\text{wb},M,g}^U} & \sqrt{D_{\text{wb},M,r}^U D_{\text{wb},M,b}^U} & \sqrt{D_{\text{wb},M,g}^U D_{\text{wb},M,b}^U} \end{bmatrix} \end{aligned} \quad (5.13)$$

3) 使用伪逆法对颜色校正矩阵的初值进行计算:

$$\begin{aligned} \mathbf{M}_0(U) &= \left( \tilde{\mathbf{D}}'_{\text{wb}}^U \right)^+ \cdot \mathbf{X}'^C \\ &= \left( \tilde{\mathbf{D}}'_{\text{wb}}^U \right)^T \cdot \left[ \tilde{\mathbf{D}}'_{\text{wb}}^U \cdot \left( \tilde{\mathbf{D}}'_{\text{wb}}^U \right)^T \right]^{-1} \cdot \mathbf{X}'^C \end{aligned} \quad (5.14)$$

4) 使用  $M$  个训练样本上的 CIEDE2000 色差均值作为目标函数对颜色校正矩阵  $\mathbf{M}(U)$  进行优化。同时, 考虑到颜色校正矩阵的中性色恒定性, 需为优化过程引入白点约束条件 (4.58)。此时, 目标函数 (4.56) 可改写为

$$\begin{aligned} \mathcal{L}_{\text{cc}} &= \frac{1}{M} \sum_{i=1}^M \Delta E_{00} \left[ \tilde{\mathbf{D}}_{\text{wb},i}^U \cdot \mathbf{M}(U), \mathbf{X}_i^C \right] \\ \text{s.t. } & \left\| \tilde{\mathbf{D}}_{n,\text{wb}} \cdot \mathbf{M}(U) - \mathbf{X}_n^C \right\|_2^2 \leq \epsilon \end{aligned} \quad (5.15)$$

其中  $\epsilon$  表示颜色校正后的白点响应值  $\tilde{\mathbf{D}}_{n,\text{wb}} \cdot \mathbf{M}(U)$  所允许的偏离参考白点  $\mathbf{X}_n^C$  的最大容限, 本论文中, 令  $\epsilon = 10^{-3}$ 。

5) 将式 (5.15) 转换为无约束优化问题:

$$\mathbf{M}(U) = \arg \min \left\{ \frac{1}{M} \sum_{i=1}^M \Delta E_{00} \left[ \tilde{\mathbf{D}}_{\text{wb},i}^U \cdot \mathbf{M}(U), \mathbf{X}_i^C \right] - \lambda \left\| \tilde{\mathbf{D}}_{n,\text{wb}} \cdot \mathbf{M}(U) - \mathbf{X}_n^C \right\|_2^2 + \lambda \epsilon \right\} \quad (5.16)$$

并使用序列二次规划算法 (sequential quadratic programming) [220] 进行求解, 得到最终的  $U$  光源对应的颜色校正矩阵  $\mathbf{M}(U)$ 。式 (5.16) 中  $\lambda$  表示拉格朗日乘数项。

表 5-1 中列出了针对本论文中所使用的两台数码相机, 在不同标定光源下利用如上非线性优化方法得到的颜色校正矩阵参数。

表 5-1 Nikon D3x 在不同标定光源下对应的二次项 RPCC 颜色校正矩阵

标定光源	D65	A	D50
颜色校正 矩阵	$\begin{bmatrix} 0.30 & 0.06 & -0.17 \\ -0.09 & 0.65 & -0.32 \\ 0.24 & -0.11 & 1.50 \\ 0.76 & 0.54 & 0.49 \\ -0.10 & -0.10 & 0.08 \\ -0.20 & -0.05 & -0.51 \end{bmatrix}$	$\begin{bmatrix} 0.21 & 0.11 & -0.23 \\ 0.09 & 0.81 & -0.17 \\ 0.25 & -0.08 & 1.80 \\ 0.66 & 0.13 & 0.48 \\ 0.04 & 0.00 & 0.41 \\ -0.34 & 0.04 & -1.20 \end{bmatrix}$	$\begin{bmatrix} 0.27 & 0.07 & -0.18 \\ -0.03 & 0.69 & -0.26 \\ 0.26 & -0.10 & 1.60 \\ 0.76 & 0.44 & 0.47 \\ -0.06 & -0.07 & 0.16 \\ -0.25 & -0.03 & -0.69 \end{bmatrix}$
标定光源	CWF	TL84	
颜色校正 矩阵	$\begin{bmatrix} 0.59 & 0.37 & -0.18 \\ 0.45 & 1.10 & -0.36 \\ 0.27 & -0.08 & 1.20 \\ -0.18 & -0.50 & 0.45 \\ 0.36 & 0.17 & 0.11 \\ -0.53 & -0.04 & -0.19 \end{bmatrix}$	$\begin{bmatrix} 0.32 & 0.15 & -0.05 \\ 0.20 & 0.47 & -0.12 \\ 0.28 & -0.12 & 1.20 \\ 0.39 & 0.27 & 0.07 \\ 0.16 & -0.12 & 0.14 \\ -0.40 & 0.35 & -0.21 \end{bmatrix}$	

### 5.3 原始响应值预测模型在图像后处理模块参数标定中的应用

#### 5.3.1 拍摄场景中白场亮度的估计

为了确定 CAT02 色适应变换模型中的适应场亮度参数  $L_A$ , 首先需要对拍摄场景中的白场亮度  $L_w$  进行估计。在场景中照明光源种类未知的情况下, 我们假设其具有等能光谱 (equi-energy spectral power distribution, 即 CIE E 光源, 指的是相对光谱功率分布函数在

各波长处恒为固定常数的假想光源)，因此只需计算该等能光谱的幅值即可完成对拍摄场景中白场亮度的估计。这一亮度估计任务的具体实施流程如下：

- 1) 从已完成非均匀性校正的图像中提取出所有像素中的最大响应值，记为  $D_{max}$ 。
- 2) 利用原始响应值预测模型对一系列具有不同幅值的等能光谱功率分布函数计算对应的原始响应值：

$$D = f[\alpha P^E(\lambda)] \quad (5.17)$$

其中  $P^E(\lambda)$  表示具有单位辐亮度的 E 光源对应的光谱功率分布函数， $\alpha$  表示等能光谱的幅值调制系数， $f(\cdot)$  代表数码相机原始响应值预测模型，其中的曝光参数（系统增益综合  $g_k$  与曝光时间  $T$ ）可从输入图像的 Exif 信息<sup>[22]</sup> 中获得。

- 3) 利用二分查找算法对最佳的幅值调制系数  $\alpha^*$  进行搜索：

$$\alpha^* = \arg \min_{\alpha} \left\{ f[\alpha P^E(\lambda)] - D_{max} \right\}^2 \quad (5.18)$$

- 4) 利用式 (2.11) 计算经  $\alpha^*$  调制的等能光谱功率分布函数  $\alpha^* P^E(\lambda)$  对应的亮度值，记为  $L_w$ ，并以此作为拍摄场景中的白场亮度。

图 5-9 展示了使用如上方法对 Nikon D3x ColorChecker 图像数据集中各测试场景进行白场亮度估计后得到的结果。

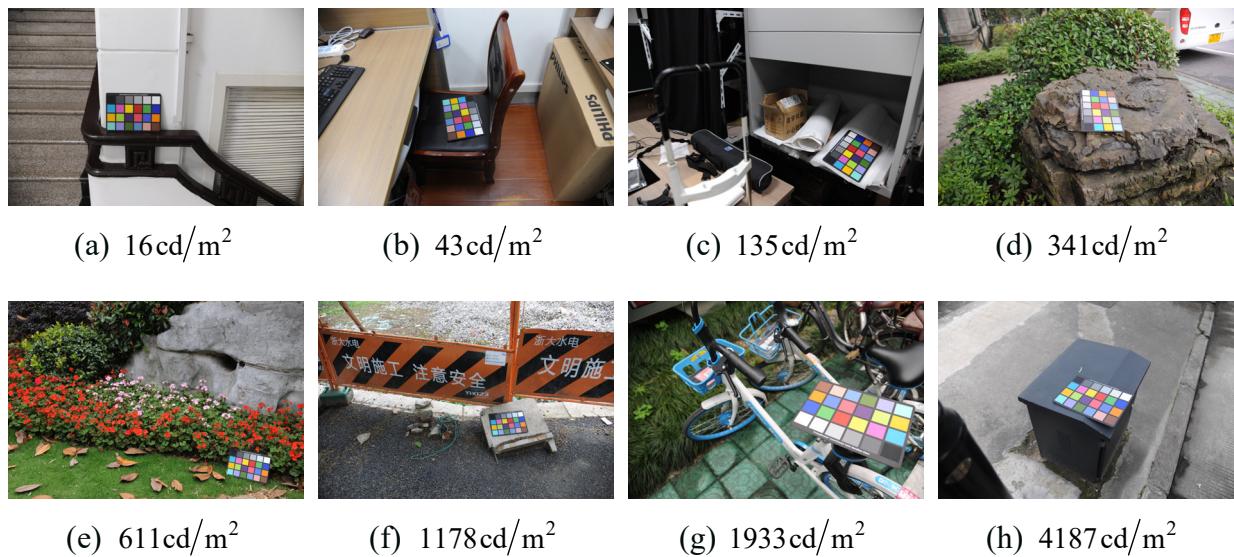


图 5-9 基于原始响应值预测模型计算得到的各场景亮度估计

## 第六章 总结与展望

### 6.1 研究工作与创新点总结

数字成像过程中的颜色复现准确性是影响图像质量的关键因素之一，也是评价数码相机性能的重要指标之一。本论文基于人眼视觉系统的感知特性，对数码相机 ISP 流程中的空间非均匀性校正、自动白平衡校正、颜色校正以及图像后处理等模块进行了原理分析与算法设计。同时，本论文提出了基于相机光谱灵敏度估计的数码相机原始响应值预测模型，该模型允许我们以仿真的手段对任意辐射体的原始响应值进行预测，从而高效地为 ISP 流程中各模块的标定实验提供大量的训练数据以及优化依据。

纵观全文，本论文的主要研究工作与创新点包括如下几个方面：

(1) 提出了完整的数码相机原始响应值构成模型。该模型充分考虑了真实系统成像过程中的光电转换特性、模数转换特性、非线性特性、噪声特性、通道串扰特性，完整描述了数码相机将空间光信号（光谱辐射亮度）转换为输出数字信号（原始数字响应值）的物理过程。同时，在该模型的基础上，提出了基于响应值构成模型的数码相机系统参数的估计方法。该方法以真实响应值与估计响应值之间的视觉感知差异作为目标函数，通过非线性优化的方法完成端到端的数码相机系统参数估计任务。结合所提出的响应值构成模型与系统参数估计方法，可对任意光源、任意拍摄对象、任意拍摄参数下的数码相机原始响应值进行准确预测，从而获得与真实相机具有几乎相同响应值特性的“虚拟相机”。

(2) 提出了基于 B-样条曲面的数字图像空间非均匀性校正算法。该算法通过对光源估计模块中得到的光源颜色信息进行分析从而自适应地选取合适的校正参数对输入图像进行空间非均匀性校正。

(3) 提出了基于像素信息统计的 GMap-NPStat 自动白平衡校正算法以及基于卷积神经网络的 ReWNet 自动白平衡校正算法。GMap-NPStat 算法综合了中性色像素统计算法与色域映射算法各自的优势，通过对中性色像素区域边界进行自适应的选择，从而改善了传统中性色统计算法无法有效应对含有大量“伪中性色”像素的图像的问题。其原理简单、直观，且无需占用过多的计算资源，因此十分适合部署于现阶段移动终端内的数码相机之中。ReWNet 算法将传统的统计类自动白平衡算法中“先手动设计约束条件，再根据条件进行像素筛选”的过程隐含于二次卷积单元（ReWU）的卷积操作之中，并通过对图像中能够反映光源颜色的特征进行自动学习从而完成光源颜色估计任务。相对于其它基于深度

学习的自动白平衡算法而言，ReWNet 算法具有更强的可解释性，且仅需占用 8%~20% 的计算资源以及 1%~3% 的参数存储开销。在公开的颜色恒常性基准数据集上，ReWNet 算法的性能与当前最好的算法（state-of-the-art）基本持平。

(4) 对数码相机颜色复现过程中的物体色同色异谱失配现象进行了定量分析。基于大量真实光谱反射比数据的仿真实验结果表明，物体色同色异谱失配并不是造成数字成像设备颜色失真的主要原因。相比于物体色同色异谱，相机光谱灵敏度函数与标准观察者色匹配函数之间的非线性关系以及光源色同色异谱失配现象，才是导致数码相机颜色复现精度出现劣化的主要原因。

(5) 提出了基于 CAT02 色适应变换模型的图像颜色外貌调节算法。该算法以人眼视觉系统真实感知特性为导向，通过对图像的各通道强度进行适当的调节从而更加真实地还原出照明光源对于场景中物体颜色外貌的渲染作用。

(6) 借助多个实例展示了原始响应值预测模型在 ISP 流程各模块调试过程中所起的作用。利用原始响应值预测模型生成的仿真数据，可实现实验室环境下的图像处理模块参数优化，由此大幅提高数码相机的 ISP 调试效率。

## 6.2 存在的不足与未来工作展望

随着数字成像技术的发展与普及，消费者对于获取更高质量数字图像的需求也在与日俱增。从目前的行业发展趋势来看，商用数码相机（尤其是移动终端内的数码相机）的设计者们愈加倾向于通过提升软件计算能力的方法以实现对数码相机图像采集能力、数据分析能力进行扩展与增强。一套实用的 ISP 流程所包含的单元与模块要远比本文中所提及的更为复杂，且往往需要在性能与资源开销之间寻求合适的妥协。由于作者水平、精力的限制，本文仅针对“ISP 流程中的颜色准确复现”这一研究目标进行了相关的探索与分析，而并未涉及与图像动态范围、清晰程度、主观喜好性等研究方向有关的讨论。

回顾全文，部分研究工作仍然存在不足之处，有待于今后进一步改进与完善。

(1) 本文所设计的 ISP 流程未涉及反马赛克 (demosaicing) 模块的相关内容。一套实用的数码相机 ISP 流程通常会采用色彩插值的方式从单通道的彩色滤色片阵列 (CFA) 输出图像中恢复出三通道 RGB 图像，而本论文仅采用最简单的降分辨率下采样的方法对 CFA 数据进行处理，并未考虑反马赛克算法对于最终的颜色复现精度可能产生的影响。

(2) 在数码相机 ISP 流程中，空间非均匀性校正模块与自动白平衡校正模块实际上

构成了一对相互掣肘的矛盾：在自动白平衡校正模块对光源颜色进行估计之前，空间非均匀性校正模块无法根据光源种类选取合适的校正参数对图像进行校正；而在空间非均匀性校正模块对输入图像进行校正之前，图像中的颜色非均匀性（color shading）又有很大的概率导致自动白平衡校正模块在对光源颜色进行估计时产生误差。在本论文的 ISP 流程设计中，首先使用参考光源（D65）对应的空间非均匀性校正参数对任意输入图像进行校正，然后在此基础上进行自动白平衡校正，最后再重复 1~2 次空间非均匀性校正+自动白平衡校正的操作。然而，这种迭代的校正策略耗费了大量的计算资源以及执行时间，因此并不是一种理想的解决方案。为解决这一“鸡与蛋的悖论”，有必要使用某些基于图像内容的、无需预先标定过程（calibration-free）的非均匀性校正算法对输入图像进行预处理，从而解除这两个模块间的相互依赖关系。

(3) 在颜色恒常性研究领域中，对大面积的单色场景进行光源颜色估计始终是令研究者最为头疼的问题之一（事实上，即使是人眼视觉系统，在面对此类语境信息缺失的孔径色（aperture color）场景时也有很大的概率出现颜色恒常性失效<sup>[222]</sup>）。本文在对 GMap-NPStat 以及 ReWNet 自动白平衡算法进行性能评估时，并未针对性地对此类大面积单色场景进行测试。然而，对于一套需部署于实际相机系统内的算法方案而言，对于此类边缘场景（corner cases）的适应能力无疑是一项能够体现其实用价值的重要指标。此外，在本论文所提出的 ISP 流程中，我们默认了拍摄场景仅由单一的均匀光源进行照明，而在实际的拍摄环境中，往往存在着多光源照明或单光源非均匀照明的情况。对于这类情况，自动白平衡校正的问题将更加复杂，此时的光源颜色估计不再是简单的单输出回归问题。

(4) 在对 ISP 输出图像的质量进行评估时，观察者的主观喜好性是一项重要的考察指标。由于时间有限，本文并未实施任何心理物理学实验对 ISP 中各模块的输出图像进行主观图像质量的评价。此外，本论文中所有模型的设计与优化均只针对“提高颜色复现精度”这一目标，若将该目标修改为“提高观察者主观喜好性”，也不失为另一个有价值的研究方向。

除了本论文中涉及的数字图像颜色复现相关研究之外，数码相机 ISP 相关领域中还存在如下一些具有探索价值与应用前景的研究方向：

(1) 当数码相机输出图像的接收对象不再是人类观察者而是机器时，数字成像技术以及 ISP 流程设计是否需要进行相应的修改？随着物联网、智能安防、无人驾驶等行业的兴起，数字成像及处理技术在越来越多的应用中扮演着对数据获取与行为决策两个环节进

行衔接的角色，此时如何根据下游计算机视觉任务的需要对 ISP 输出图像进行调制与增强，也将成为一个值得探究的课题。

(2) 随着深度学习以及强化学习技术的发展，在未来是否有可能实现端到端的 ISP 解决方案？关于这一问题，Jiang 等人<sup>[223]</sup>、Chen 等人<sup>[224]</sup>以及 Schwartz 等人<sup>[225]</sup>在近期均进行了一些探索性的尝试，然而此类方案或多或少存在着可解释性差、难以预留调试接口、标定过程耗时长等问题，在正式投入产品应用之前仍然面临着不少挑战。因此，如何对 ISP 中的各个模块进行组织与整合以实现全局的图像质量优化，也将成为该领域内一个有意义的研究方向。

(3) 在未来可预期的一段时间内，多光谱甚至高光谱成像技术是否有可能像目前的三色成像技术一样得到普及？正如前文所述的那样，由于光谱信息的缺失，三色相机在颜色复现的应用上存在一些本质的缺陷，因此，如何令数字成像系统兼具三色相机与光谱相机各自的优势，并将目前仅针对专业领域的光谱成像技术推广至更多的生活生产应用领域之中，必定将成为该领域内的一个热门研究方向。

## 参考文献

- [1] Stoffel J C, Moreland J. A survey of electronic techniques for pictorial image reproduction[J]. IEEE Transactions on Communications, 1981, 29: 1898–1925.
- [2] Marion B, Bernard C A. Image reproduction processes: 3060023[P]. October 1962.
- [3] Cox G G, Cook L T, Mcmillan J H, et al. Chest radiography: comparison of high-resolution digital displays with conventional and digital film.[J]. Radiology, 1990, 176: 771–776.
- [4] Reinhard E, Heidrich W, Debevec P, et al. High dynamic range imaging: acquisition, display, and image-based lighting[M]. Morgan Kaufmann, 2010.
- [5] Buades A, Coll B, Morel J-M. A review of image denoising algorithms, with a new one[J]. Multiscale Modeling & Simulation, 2005, 4(2): 490-530.
- [6] Yamaguchi M, Teraji T, Ohsawa K, et al. Color image reproduction based on multispectral and multiprimary imaging: experimental evaluation[C]. Color Imaging: Device-Independent Color, Color Hardcopy, and Applications VII, 2001: 15-27.
- [7] Faugeras O. Digital color image processing within the framework of a human visual model[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1979, 27(4): 380-393.
- [8] Huang K-Q, Wang Q, Wu Z-Y. Natural color image enhancement and evaluation algorithm based on human visual system[J]. Computer Vision and Image Understanding, 2006, 103(1): 52-63.
- [9] Gindele E B, Topfer K, Buhr J D, et al. Color image reproduction of scenes with preferential color mapping and scene-dependent tone scaling: 6594388[P]. July 2003.
- [10] Rahman Z-U, Jobson D J, Woodell G A. Retinex processing for automatic image enhancement[J]. Journal of Electronic Imaging, 2004, 13(1): 100-111.
- [11] Hordley S D. Scene illuminant estimation: past, present, and future[J]. Color Research & Application, 2006, 31(4): 303-314.
- [12] Qiu J, Xu H, Ye Z, et al. Image quality degradation of object-color metamer mismatching in digital camera color reproduction[J]. Applied Optics, 2018, 57(11): 2851-2860.
- [13] Taigman Y, Yang M, Ranzato M A, et al. Deepface: Closing the gap to human-level performance in face verification[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1701-1708.
- [14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [15] Rui Y, Huang T S, Chang S-F. Image retrieval: Current techniques, promising directions, and

- open issues[J]. Journal of visual communication and image representation, 1999, 10(1): 39-62.
- [16] Kerl C, Souiai M, Sturm J, et al. Towards illumination-invariant 3d reconstruction using TOF RGB-D cameras[C]. International Conference on 3D Vision, 2014: 39-46.
- [17] Murthy J K, Krishna G, Chhaya F, et al. Reconstructing vehicles from a single image: Shape priors for road scene understanding[J]. arXiv preprint arXiv:1609.09468, 2016.
- [18] Katsaggelos A K. Digital image restoration[M]. Springer Publishing Company, Incorporated, 2012.
- [19] Wang S, Zhang Y, Deng P, et al. Fast automatic white balancing method by color histogram stretching[C]. International Congress on Image and Signal Processing, 2011: 979-983.
- [20] Mauer C. Measurement of the spectral response of digital cameras with a set of interference filters[J]. Thesis at the Department of Media-and Phototechnology, 2009.
- [21] Hubel P M, Sherman D, Farrell J E. A comparison of methods of sensor spectral sensitivity estimation[C]. Color and Imaging Conference, 1994: 45-48.
- [22] Barnard K, Funt B. Camera characterization for color research[J]. Color Research & Application, 2002, 27(3): 152-163.
- [23] Jiang J, Liu D, Gu J, et al. What is the space of spectral sensitivity functions for digital color cameras?[C]. IEEE Workshop on Applications of Computer Vision, 2013: 168–179.
- [24] Prasad D, Nguyen R, Brown M. Quick Approximation of Camera's Spectral Response from Casual Lighting[C]. IEEE International Conference on Computer Vision Workshops, 2013: 844-851.
- [25] Darrodi M M, Finlayson G, Goodman T, et al. Reference data set for camera spectral sensitivity estimation[J]. Journal of the Optical Society of America A, 2015, 32(3): 381-391.
- [26] Martínez-Verdú F, Pujol J, Capilla P. Characterization of a digital camera as an absolute tristimulus colorimeter[J]. Journal of Imaging Science and Technology, 2003, 47(4): 279-295.
- [27] 张显斗. 数字图像颜色复现理论与方法研究[D]. 光电科学与工程学院, 浙江大学, 2010.
- [28] Grossberg M D, Nayar S K. What is the space of camera response functions?[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2003: 602.
- [29] Grossberg M D, Nayar S K. Determining the camera response from images: What is knowable?[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(11): 1455-1467.

- [30] Hong G, Luo M R, Rhodes P A. A study of digital camera colorimetric characterization based on polynomial modeling[J]. *Color Research & Application*, 2001, 26(1): 76-84.
- [31] Bianco S, Schettini R, Vanneschi L. Empirical modeling for colorimetric characterization of digital cameras[C]. *IEEE International Conference on Image Processing*, 2009: 3469-3472.
- [32] Hardeberg J Y. Colorimetric characterization of digital cameras preserving hue planes[C]. *Color and Imaging Conference*, 2005: 141-146.
- [33] Liu Y, Yu H, Shi J. Camera characterization using back-propagation artificial neural network based on Munsell system[C]. *International Symposium on Photoelectronic Detection and Imaging 2007: Photoelectronic Imaging and Detection*, 2008: 66210A.
- [34] Münzenmayer C, Naujokat F, Mühlendorfer S, et al. Enhancing texture analysis by color shading correction[C]. *Workshop Farbbildverarbeitung*, 2003.
- [35] Young I T. Shading correction: compensation for illumination and sensor inhomogeneities[J]. *Current Protocols in Cytometry*, 2000, 14(1): 2.11. 1-2.11. 12.
- [36] Yu W. Practical anti-vignetting methods for digital cameras[J]. *IEEE Transactions on Consumer Electronics*, 2004, 50(4): 975-983.
- [37] Tajbakhsh T. Color lens shade compensation achieved by linear regression of piece-wise bilinear spline functions[C]. *Digital Photography*, 2010: 75370P.
- [38] Likar B, Maintz J A, Viergever M A, et al. Retrospective shading correction based on entropy minimization[J]. *Journal of Microscopy*, 2000, 197(Pt 3): 285-295.
- [39] Lopez-Fuentes L, Oliver G, Massanet S. Revisiting image vignetting correction by constrained minimization of log-intensity entropy[C]. *International Work-Conference on Artificial Neural Networks*, 2015: 450-463.
- [40] Zheng Y, Lin S, Kang S B, et al. Single-image vignetting correction from gradient distribution symmetries[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6): 1480-1494.
- [41] Lam E Y. Combining gray world and retinex theory for automatic white balance in digital photography[C]. *International Symposium on Consumer Electronics*, 2005: 134–139.
- [42] Rizzi A, Gatta C, Marini D. Color correction between gray world and white patch[C]. *Human Vision and Electronic Imaging*, 2002: 367–375.
- [43] Van De Weijer J, Gevers T, Gijsenij A. Edge-based color constancy[J]. *IEEE Transactions on image processing*, 2007, 16(9): 2207-2214.
- [44] Gijsenij A, Gevers T, Van De Weijer J. Physics-based edge evaluation for improved color constancy[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 581-

588.

- [45] Van De Weijer J, Gevers T. Color constancy based on the grey-edge hypothesis[C]. IEEE International Conference on Image Processing, 2005: II-722.
- [46] Chiu L-C, Fuh C-S. Calibration-based auto white balance method for digital still camera[J]. Journal of Information Science and Engineering, 2010, 26(2).
- [47] Chikane V, Fuh C-S. Automatic white balance for digital still cameras[J]. Journal of Information Science and Engineering, 2006, 22(3): 497-509.
- [48] Weng C-C, Chen H, Fuh C-S. A novel automatic white balance method for digital still cameras[C]. Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, 2005: 3801-3804.
- [49] Huo J-Y, Chang Y-L, Wang J, et al. Robust automatic white balance algorithm using gray color points in images[J]. IEEE Transactions on Consumer Electronics, 2006, 52(2): 541-546.
- [50] Gao S, Han W, Yang K, et al. Efficient color constancy with local surface reflectance statistics[C]. European Conference on Computer Vision, 2014: 158-173.
- [51] Forsyth D A. A novel algorithm for color constancy[J]. International Journal of Computer Vision, 1990, 5(1): 5-35.
- [52] Finlayson G D, Hordley S D, Tastl I. Gamut constrained illuminant estimation[J]. International Journal of Computer Vision, 2006, 67(1): 93-109.
- [53] Gijsenij A, Gevers T, Van De Weijer J. Generalized gamut mapping using image derivative structures for color constancy[J]. International Journal of Computer Vision, 2010, 86(2): 127-139.
- [54] Finlayson G D, Hordley S D, Hubel P M. Color by correlation: A simple, unifying framework for color constancy[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(11): 1209-1221.
- [55] Kehtarnavaz N, Kim N, Gamadia M. Real-time auto white balancing using DWT-based multi-scale clustering[C]. Real-Time Image Processing, 2006: 606308.
- [56] Funt B, Xiong W. Estimating illumination chromaticity via support vector regression[C]. Color and Imaging Conference, 2004: 47-52.
- [57] Gehler P V, Rother C, Blake A, et al. Bayesian color constancy revisited[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2008: 1-8.
- [58] Cheng D, Price B, Cohen S, et al. Effective learning-based illuminant estimation using simple features[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015:

- 1000-1008.
- [59] Bianco S, Cusano C, Schettini R. Color constancy using CNNs[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 81-89.
- [60] Mcdonagh S, Parisot S, Li Z, et al. Meta-learning for few-shot camera-adaptive color constancy[J]. arXiv preprint arXiv:1811.11788, 2018.
- [61] Zapryanov G, Ivanova D, Nikolova I. Automatic white balance algorithms for digital still cameras—a comparative study[C]. IEEE International Conference on Multimedia and Expo, 2012.
- [62] Finlayson G D, Drew M S. White-point preserving color correction[C]. Color and Imaging Conference, 1997: 258-261.
- [63] Finlayson G D, Mackiewicz M, Hurlbert A. Color correction using root-polynomial regression[J]. IEEE Transactions on image processing, 2015, 24(5): 1460-1470.
- [64] Cherdhirunkorn K, Tsumura N, Nakaguchi T, et al. Spectral based color correction technique compatible with standard RGB system[J]. Optical Review, 2006, 13(3): 138-145.
- [65] Marguier J, Bhatti N, Baker H, et al. Color correction of uncalibrated images for the classification of human skin color[C]. Color and Imaging Conference, 2007: 331-335.
- [66] Zhang H, Liu H, Quan S. Hue constrained matrix optimization for preferred color reproduction[J]. Journal of Electronic Imaging, 2012, 21(3): 033021-1-033021-15.
- [67] 徐海松. 颜色信息工程[M]. 第二版. 浙江大学出版社, 2015.
- [68] Wyszecki G, Stiles W S. Color science[M]. 8. Wiley New York, 1982.
- [69] Prasad D K, Wenhe L. Metrics and statistics of frequency of occurrence of metamerism in consumer cameras for natural scenes[J]. Journal of the Optical Society of America A, 2015, 32(7): 1390-1402.
- [70] Nimeroff I, Yurow J. Degree of metamerism[J]. Journal of the Optical Society of America, 1965, 55(2): 185-190.
- [71] Logvinenko A D, Funt B, Mirzaei H, et al. Rethinking colour constancy[J]. PloS one, 2015, 10(9): e0135029.
- [72] Zhang X, Funt B, Mirzaei H. Metamer mismatching in practice versus theory[J]. Journal of the Optical Society of America A, 2016, 33(3): A238-A247.
- [73] Foster D, Marín-Franch I. Effectiveness of digital camera sensors in distinguishing colored surfaces in outdoor scenes[C]. Imaging Systems and Applications, 2013: ITh3D. 2.
- [74] Sanfeliu A, Cortés M L. Progress in pattern recognition, image analysis and applications[C]. Iberoamerican Congress on Pattern Recognition, 2004.

- [75] Fairchild M D, Rosen M R, Johnson G M. Spectral and metameric color imaging[J]. RIT-MCSL Technical Report, 2001.
- [76] Zhang X, Funt B, Mirzaei H. Metamer mismatching and its consequences for predicting how colours are affected by the illuminant[C]. IEEE International Conference on Computer Vision Workshops, 2015: 1-7.
- [77] Hung P-C. Sensitivity metamerism index for digital still camera[C]. Proc. SPIE, 2002.
- [78] Iso. Graphic technology and photography -- Colour characterisation of digital still cameras (DSCs) -- Part 1: Stimuli, metrology and test procedures[S]. 2012.
- [79] 叶关荣. 光辐射测量[M]. 浙江大学光电系内部教材, 不详.
- [80] Stiles W, Burch J. Interim report to the Commission Internationale de l'Eclairage, Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching (1955)[J]. Optica Acta: International Journal of Optics, 1955, 2(4): 168-181.
- [81] Grassmann H. Zur theorie der farbenmischung[J]. Annalen der Physik, 1853, 165(5): 69-84.
- [82] Wright W D. A re-determination of the trichromatic coefficients of the spectral colours[J]. Transactions of the Optical Society, 1929, 30(4): 141.
- [83] Guild J. A trichromatic colorimeter suitable for standardisation work[J]. Transactions of the Optical Society, 1925, 27(2): 106.
- [84] Hunt R W. The heights of the CIE colour - matching functions[J]. Color Research & Application, 1997, 22(5): 335-335.
- [85] Fairman H S, Brill M H, Hemmendinger H. How the CIE 1931 color - matching functions were derived from Wright - Guild data[J]. Color Research & Application, 1997, 22(1): 11-23.
- [86] 方竞宇. 高动态范围彩色图像捕获与显示方法及技术研究[D]. 光电科学与工程学院, 浙江大学, 2016.
- [87] Stiles W S, Burch J M. NPL colour-matching investigation: final report (1958)[J]. Optica Acta: International Journal of Optics, 1959, 6(1): 1-26.
- [88] Speranskaya N. Determination of spectrum color coordinates for twenty-seven normal observers[J]. Opt. Spectrosc., 1959, 7: 424-428.
- [89] 吴晓玟. 基于高动态范围成像的光环境测量研究[D]. 光电科学与工程学院, 浙江大学, 2017.
- [90] Sharma G, Wu W, Dalal E N. The CIEDE2000 color - difference formula: Implementation notes, supplementary test data, and mathematical observations[J]. Color Research &

- Application, 2005, 30(1): 21-30.
- [91] Planck M. The theory of heat radiation[M]. Courier Corporation, 2013.
- [92] Ohta N, Robertson A. Colorimetry: fundamentals and applications[M]. John Wiley & Sons, 2006.
- [93] Fairchild M D. Color appearance models[M]. John Wiley & Sons, 2013.
- [94] Nayatani Y, Takahama K, Sobagaki H, et al. Color - appearance model and chromatic - adaptation transform[J]. Color Research & Application, 1990, 15(4): 210-221.
- [95] Hunt R. Revised colour - appearance model for related and unrelated colours[J]. Color Research & Application, 1991, 16(3): 146-165.
- [96] Fairchild M D. Refinement of the RLAB color space[J]. Color Research & Application, 1996, 21(5): 338-346.
- [97] Cie Tc8-01. A colour appearance model for colour management systems: CIECAM02[R]. Vienna: CIE, 2004.
- [98] Horn B K, Sjoberg R W. Calculating the reflectance map[J]. Applied Optics, 1979, 18(11): 1770-1779.
- [99] Walree P V. Vignetting[EB/OL]. <http://www.cs.cmu.edu/~sensing-sensors/readings/vignetting.pdf>.
- [100] Martin C B. Design issues of a hyperfield fisheye lens[C]. Novel Optical Systems Design and Optimization, 2004: 84-93.
- [101] Aggarwal M, Hua H, Ahuja N. On cosine-fourth and vignetting effects in real lenses[C]. IEEE International Conference on Computer Vision, 2001: 472-479.
- [102] 郁道银, 谈恒英. 工程光学[M]. 北京: 机械工业出版社, 1999.
- [103] Smith W J. Modern optical engineering[M]. Tata McGraw-Hill Education, 1966.
- [104] Cauwerts C, Phd M B, Deneyer A. Comparison of the vignetting effects of two identical fisheye lenses[J]. Leukos, 2012, 8(3): 181-203.
- [105] Yamamoto K. Image sensor having large micro-lenses at the peripheral regions. Google Patents, 2003.
- [106] Bayer B E. Color imaging array. Google Patents, 1976.
- [107] Wikipedia Contributors. Fujifilm X-Trans sensor[EB/OL]. (14 February 2019 20:11 UTC)[8 March 2019 16:47 UTC]. [https://en.wikipedia.org/w/index.php?title=Fujifilm\\_X-Trans\\_sensor&oldid=883337483](https://en.wikipedia.org/w/index.php?title=Fujifilm_X-Trans_sensor&oldid=883337483).
- [108] Wikipedia Contributors. Foveon X3 sensor[EB/OL]. (1 October 2018 20:35 UTC)[8 March 2019 16:54 UTC]. [https://en.wikipedia.org/w/index.php?title=Foveon\\_X3\\_sensor&oldid=883337483](https://en.wikipedia.org/w/index.php?title=Foveon_X3_sensor&oldid=883337483).

- 862046984.
- [109] Healey G E, Kondepudy R. Radiometric CCD camera calibration and noise estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(3): 267-276.
- [110] Baker R J. CMOS: circuit design, layout, and simulation[M]. 1. John Wiley & Sons, 2008.
- [111] Cao Y, Pan X, Zhao X, et al. An analog gamma correction scheme for high dynamic range CMOS logarithmic image sensors[J]. Sensors, 2014, 14(12): 24132-24145.
- [112] Ham S, Lee Y, Jung W, et al. CMOS image sensor with analog gamma correction using nonlinear single-slope ADC[C]. IEEE International Symposium on Circuits and Systems, 2006: 4 pp.-3581.
- [113] European Machine Vision Association. EMVA Standard 1288, Standard for Characterization of Image Sensors and Cameras[S]. Release, 2010: 29.
- [114] Steger C, Ulrich M, Wiedemann C. Machine vision algorithms and applications[M]. John Wiley & Sons, 2018.
- [115] Wronski B, Garcia-Dorado I, Ernst M, et al. Handheld multi-frame super-resolution[J]. arXiv preprint arXiv:1905.03277, 2019.
- [116] Martinec E. Noise, dynamic range and bit depth in digital SLRs[EB/OL]. <http://theory.uchicago.edu/~ejm/pix/20d/tests/noise/index.html>.
- [117] Barry J R, Lee E A, Messerschmitt D G. Digital communication[M]. Springer Science & Business Media, 2012.
- [118] Perepelitsa D V. Johnson noise and shot noise[EB/OL]. <http://web.mit.edu/dvp/Public/noise-paper.pdf>.
- [119] Grant L. Characterisation of noise sources in CMOS image sensors[C]. IEEE Solid-State Circuits Conf. Image Sensor Design Forum, 2005.
- [120] Lim S, Silverstein A. Spatially varying color correction (SVCC) matrices for reduced noise[C]. Color and Imaging Conference, 2004: 76-81.
- [121] Mildenhall B, Barron J T, Chen J, et al. Burst denoising with kernel prediction networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2502-2510.
- [122] Zhang F, Xu H, Wang Z. Spectral design methods for multi-channel LED light sources based on differential evolution[J]. Applied Optics, 2016, 55(28): 7771-7781.
- [123] McCullagh P. Generalized linear models[M]. Routledge, 2018.
- [124] Myers R H, Myers R H. Classical and modern regression with applications[M]. 2. Duxbury Press Belmont, CA, 1990.
- [125] Agranov G, Berezin V, Tsai R H. Crosstalk and microlens study in a color CMOS image

- sensor[J]. IEEE Transactions on Electron Devices, 2003, 50(1): 4-11.
- [126] Chan W W, Sin J K, Wong S S. A novel crosstalk isolation structure for bulk CMOS power IC's[J]. IEEE Transactions on Electron Devices, 1998, 45(7): 1580-1586.
- [127] Urban P, Desch M, Happel K, et al. Recovering camera sensitivities using target-based reflectances captured under multiple LED-illuminations[C]. Workshop on Color Image Processing, 2010: 9-16.
- [128] Finlayson G D, Hordley S, Hubel P M. Recovering device sensitivities with quadratic programming[C]. Color and Imaging Conference, 1998: 90-95.
- [129] Huynh C P, Robles-Kelly A. Recovery of spectral sensitivity functions from a colour chart image under unknown spectrally smooth illumination[C]. International Conference on Pattern Recognition, 2014: 708-713.
- [130] Ebner M. Estimating the spectral sensitivity of a digital sensor using calibration targets[C]. Annual Conference on Genetic and Evolutionary Computation, 2007: 642-649.
- [131] Byrd R H, Hribar M E, Nocedal J. An interior point algorithm for large-scale nonlinear programming[J]. SIAM Journal on Optimization, 1999, 9(4): 877-900.
- [132] Forsgren A, Gill P E, Wright M H. Interior methods for nonlinear optimization[J]. SIAM review, 2002, 44(4): 525-597.
- [133] Hansen P C. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems[J]. Numerical algorithms, 1994, 6(1): 1-35.
- [134] Hansen P C, O'leary D P. The use of the L-curve in the regularization of discrete ill-posed problems[J]. SIAM Journal on Scientific Computing, 1993, 14(6): 1487-1503.
- [135] Lawson C L, Hanson R J. Solving least squares problems[M]. 15. Siam, 1995.
- [136] Mathworks. lsqnonneg - Solve nonnegative linear least-squares problem[EB/OL]. <https://www.mathworks.com/help/matlab/ref/lsqnonneg.html>.
- [137] Coleman T F, Y L. An interior trust region approach for nonlinear minimization subject to bounds[J]. SIAM Journal on Optimization, 1996, 6(2): 418-445.
- [138] Zhao H, Kawakami R, Tan R T, et al. Estimating basis functions for spectral sensitivity of digital cameras[C]. Meeting on Image Recognition and Understanding, 2009.
- [139] Sharma G, Bala R. Digital color imaging handbook[M]. CRC press, 2002.
- [140] 薛江. 图像信号处理器 (ISP) 的实现和 FPGA 验证[D]. 浙江大学, 2007.
- [141] Lukac R. Computational photography: methods and applications[M]. CRC Press, 2016.
- [142] Photography D D. What is Vignetting? How to remove or add vignetting to a photo[EB/OL]. <https://www.discoverdigitalphotography.com/2013/what-is-vignetting-how-to-remove-or->

[add-vignetting-to-a-photo/](#).

- [143] Phillips J B, Eliasson H. Camera image quality benchmarking[M]. John Wiley & Sons, 2018.
- [144] Peltoketo V-T. Benchmarking of mobile phone cameras[D]. Department of computer science, University of Vaasa, 2016.
- [145] Goldman D B. Vignette and exposure calibration and compensation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(12): 2276-2288.
- [146] Eck M, Hoppe H. Automatic reconstruction of B-spline surfaces of arbitrary topological type[M]. Fachbereich Mathematik, Technische Hochschule Darmstadt, 1996.
- [147] Eberly D. Least-squares fitting of data with B-spline surfaces[EB/OL]. <https://www.geometrytools.com/Documentation/BSplineSurfaceLeastSquaresFit.pdf>.
- [148] Geusebroek J-M, Van Den Boomgaard R, Smeulders A W, et al. Color constancy from physical principles[J]. Pattern Recognition Letters, 2003, 24(11): 1653-1662.
- [149] Chittka L, Menzel R. The evolutionary adaptation of flower colours and the insect pollinators' colour vision[J]. Journal of Comparative Physiology A, 1992, 171(2): 171-181.
- [150] Zhang X-S, Gao S-B, Li R-X, et al. A retinal mechanism inspired color constancy model[J]. IEEE Transactions on image processing, 2016, 25(3): 1219-1232.
- [151] Thompson E. Colour vision, evolution, and perceptual content[J]. Synthese, 1995, 104(1): 1-32.
- [152] Delahunt P B, Brainard D H. Does human color constancy incorporate the statistical regularity of natural daylight?[J]. Journal of Vision, 2004, 4(2): 1-1.
- [153] Witzel C, Gegenfurtner K R. Color perception: Objects, constancy, and categories[J]. Annual review of vision science, 2018, 4: 475-499.
- [154] Hurlbert A C, Ling Y. If it's a banana, it must be yellow: The role of memory colors in color constancy[J]. Journal of Vision, 2005, 5(8): 787-787.
- [155] Jin E W, Shevell S K. Color memory and color constancy[J]. Journal of the Optical Society of America A, 1996, 13(10): 1981-1991.
- [156] Rahtu E, Nikkanen J, Kannala J, et al. Applying visual object categorization and memory colors for automatic color constancy[C]. International Conference on Image Analysis and Processing, 2009: 873-882.
- [157] Liu Y-C, Chan W-H, Chen Y-Q. Automatic white balance for digital still camera[J]. IEEE Transactions on Consumer Electronics, 1995, 41(3): 460-466.
- [158] Gijsenij A, Gevers T, Van De Weijer J. Computational color constancy: Survey and

- experiments[J]. IEEE Transactions on image processing, 2011, 20(9): 2475-2489.
- [159] Ebner M. Color constancy[M]. 7. John Wiley & Sons, 2007.
- [160] Finlayson G, Hordley S. Improving gamut mapping color constancy[J]. IEEE Transactions on image processing, 2000, 9(10): 1774-1783.
- [161] Fryc I, Brown S W, Ohno Y. Spectral matching with an LED-based spectrally tunable light source[C]. International Conference on Solid State Lighting, 2005: 5941I.
- [162] Tominaga S, Ebisui S, Wandell B A. Scene illuminant classification: brighter is better[J]. Journal of the Optical Society of America A, 2001, 18(1): 55-64.
- [163] Yang K-F, Gao S-B, Li Y-J. Efficient illuminant estimation for color constancy using grey pixels[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2254-2263.
- [164] Zakizadeh R, Brown M S, Finlayson G D. A hybrid strategy for illuminant estimation targeting hard images[C]. IEEE International Conference on Computer Vision Workshops, 2015: 16-23.
- [165] Brainard D H, Wandell B A. Analysis of the retinex theory of color vision[J]. Journal of the Optical Society of America A, 1986, 3(10): 1651-1661.
- [166] Buchsbaum G. A spatial processor model for object colour perception[J]. Journal of the Franklin institute, 1980, 310(1): 1-26.
- [167] Finlayson G D, Trezzi E. Shades of gray and colour constancy[C]. Color and Imaging Conference, 2004: 37-41.
- [168] Coffin D J. Decoding raw digital photos in Linux[EB/OL]. 2019.05.20]. <https://www.cybercom.net/~dc coffin/dcraw/>.
- [169] Anderson M, Motta R, Chandrasekar S, et al. Proposal for a standard default color space for the internet—sRGB[C]. Color and Imaging Conference, 1996: 238-245.
- [170] Qian Y, Chen K, Kämäärinen J-K, et al. Deep structured-output regression learning for computational color constancy[C]. International Conference on Pattern Recognition, 2016: 1899-1904.
- [171] Cardei V C, Funt B, Barnard K. Estimating the scene illumination chromaticity by using a neural network[J]. Journal of the Optical Society of America A, 2002, 19(12): 2374-2386.
- [172] Chakrabarti A. Color constancy by learning to predict chromaticity from luminance[C]. Advances in Neural Information Processing Systems, 2015: 163-171.
- [173] Barron J T. Convolutional color constancy[C]. IEEE International Conference on Computer Vision, 2015: 379-387.
- [174] Barron J T, Tsai Y-T. Fast fourier color constancy[C]. IEEE Conference on Computer Vision

- and Pattern Recognition, 2017: 886-894.
- [175] Hu Y, Wang B, Lin S. FC4: fully convolutional color constancy with confidence-weighted pooling[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4085-4094.
- [176] Hold-Geoffroy Y, Sunkavalli K, Hadap S, et al. Deep outdoor illumination estimation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7312-7321.
- [177] Afifi M. Semantic white balance: Semantic color constancy using convolutional neural network[J]. arXiv preprint arXiv:1802.00153, 2018.
- [178] Shi W, Loy C C, Tang X. Deep specialized network for illuminant estimation[C]. European Conference on Computer Vision, 2016: 371-387.
- [179] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [180] Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[C]. International Conference on Machine Learning, 2010: 807-814.
- [181] Lin M, Chen Q, Yan S. Network in network[J]. arXiv preprint arXiv:1312.4400, 2013.
- [182] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.
- [183] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint arXiv:1502.03167, 2015.
- [184] Lei Ba J, Kiros J R, Hinton G E. Layer normalization[J]. arXiv preprint arXiv:1607.06450, 2016.
- [185] Ulyanov D, Vedaldi A, Lempitsky V. Instance normalization: The missing ingredient for fast stylization[J]. arXiv preprint arXiv:1607.08022, 2016.
- [186] Hemrit G, Finlayson G D, Gijsenij A, et al. Rehabilitating the ColorChecker dataset for illuminant estimation[C]. Color and Imaging Conference, 2018: 350-353.
- [187] Cheng D, Prasad D K, Brown M S. Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution[J]. Journal of the Optical Society of America A, 2014, 31(5): 1049-1058.
- [188] Shi L. Re-processed version of the gehler color constancy dataset of 568 images[EB/OL].  
<http://www.cs.sfu.ca/~color/data/>.
- [189] Banić N, Lončarić S. Unsupervised learning for color constancy[J]. arXiv preprint arXiv:1712.00436, 2017.

- [190] American National Standards Institute. Specifications for the chromaticity of solid state lighting products[S]. 2011.
- [191] Abadi M, Barham P, Chen J, et al. Tensorflow: A system for large-scale machine learning[C]. Symposium on Operating Systems Design and Implementation, 2016: 265-283.
- [192] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [193] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [194] Barnard K. Improvements to gamut mapping colour constancy algorithms[C]. European Conference on Computer Vision, 2000: 390-403.
- [195] Chakrabarti A, Hirakawa K, Zickler T. Color constancy with spatio-spectral statistics[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(8): 1509-1519.
- [196] Schanda J. Colorimetry: understanding the CIE system[M]. John Wiley & Sons, 2007.
- [197] Mokrzycki W, Tatol M. Colour difference  $\Delta E$ -a survey[J]. Machine Graphics and Vision, 2011, 20(4): 383-411.
- [198] Mohammadi M, Nezamabadi M, Berns R S, et al. Spectral imaging target development based on hierarchical cluster analysis[C]. Color and Imaging Conference, 2004: 59-64.
- [199] Iso. Standard object colour spectra database for colour reproduction evaluation (SOCS)[S]. 2003.
- [200] Clark R N, Swayze G A, Wise R, et al. USGS digital spectral library splib06a[R]. US geological survey, digital data series, 2007: 2007.
- [201] Barnard K, Martin L, Funt B, et al. A data set for color research[J]. Color Research & Application, 2002, 27(3): 147-151.
- [202] Foster D H, Amano K, Nascimento S M, et al. Frequency of metamerism in natural scenes[J]. Journal of the Optical Society of America A, 2006, 23(10): 2359-2372.
- [203] Yasuma F, Mitsunaga T, Iso D, et al. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum[J]. IEEE Transactions on image processing, 2010, 19(9): 2241-2253.
- [204] Spectral Color Research Group. University of Eastern Finland Spectral Database[EB/OL]. <http://www.uef.fi/web/spectral/-spectral-database>.
- [205] Skauli T, Farrell J. A collection of hyperspectral images for imaging systems research[C]. Digital Photography, 2013: 86600C.

- [206] Le Moan S, George S T, Pedersen M, et al. A database for spectral image quality[C]. Image Quality and System Performance, 2015: 93960P.
- [207] Hordley S, Finalyson G, Morovic P. A multi-spectral image database and its application to image rendering across illumination[C]. International Conference on Image and Graphics, 2004: 394-397.
- [208] Párraga C, Brelstaff G, Troscianko T, et al. Color and luminance information in natural scenes[J]. Journal of the Optical Society of America A, 1998, 15(3): 563-569.
- [209] Eckhard J, Eckhard T, Valero E M, et al. Outdoor scene reflectance measurements using a Bragg-grating-based hyperspectral imager[J]. Applied Optics, 2015, 54(13): D15-D24.
- [210] Judd D B, Macadam D L, Wyszecki G, et al. Spectral distribution of typical daylight as a function of correlated color temperature[J]. Journal of the Optical Society of America, 1964, 54(8): 1031-1040.
- [211] Lee D, Plataniotis K N. Towards a full-reference quality assessment for color images using directional statistics[J]. IEEE Transactions on image processing, 2015, 24(11): 3950-3965.
- [212] Tomasi C, Manduchi R. Bilateral filtering for gray and color images[C]. IEEE International Conference on Computer Vision, 1998: 2.
- [213] Bai Y, Zhang X. Transformations and white point constraint solutions for a novel chromaticity space. Google Patents, 2015.
- [214] Parkkinen J P, Hallikainen J, Jaaskelainen T. Characteristic spectra of Munsell colors[J]. Journal of the Optical Society of America A, 1989, 6(2): 318-322.
- [215] Ansi. Graphic technology - Input data for characterization of 4-color process printing[S]. IT8.7/3, 2010.
- [216] Jaaskelainen T, Silvennoinen R, Hiltunen J, et al. Classification of the reflectance spectra of pine, spruce, and birch[J]. Applied Optics, 1994, 33(12): 2356-2362.
- [217] Haanpalo J, Silfsten P, Parkkinen J P, et al. Spectral representation of paper opacity[C]. Intelligent Robots and Computer Vision, 1997: 339-348.
- [218] Vrhel M J, Gershon R, Iwan L S. Measurement and analysis of object reflectance spectra[J]. Color Research & Application, 1994, 19(1): 4-9.
- [219] Parkkinen J, Jaaskelainen T, Kuittinen M. Spectral representation of color images[C]. International Conference on Pattern Recognition, 1988: 933-935.
- [220] Nocedal J, Wright S. Numerical optimization[M]. Springer Science & Business Media, 2006.
- [221] Wikipedia Contributors. Exif[EB/OL]. (6 June 2019 22:34 UTC)[20 June 2019 01:30 UTC].

- [https://en.wikipedia.org/w/index.php?title=Exif&oldid=900659490.](https://en.wikipedia.org/w/index.php?title=Exif&oldid=900659490)
- [222] Mccann J. Rules for colour constancy[J]. Ophthalmic and Physiological Optics, 1992, 12(2): 175-177.
- [223] Jiang H, Tian Q, Farrell J, et al. Learning the image processing pipeline[J]. IEEE Transactions on image processing, 2017, 26(10): 5032-5042.
- [224] Chen C, Chen Q, Xu J, et al. Learning to see in the dark[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3291-3300.
- [225] Schwartz E, Giryes R, Bronstein A M. DeepISP: Toward learning an end-to-end image processing pipeline[J]. IEEE Transactions on image processing, 2018, 28(2): 912-923.



## 附录一

### 关于 P17 第 2.1.1 小节公式 (2.6) 的证明

在球坐标系中，球面上面积微元所张的立体角微元  $d\omega$  可表示为

$$d\omega = \frac{r d\theta r \sin \theta d\varphi}{r^2} = \sin \theta d\theta d\varphi \quad (\text{A1.1})$$

将该式代入式 (2.6) 中第一个等号的右侧，并在上半球面  $\Omega = 2\pi \text{ sr}$  的立体角（对应  $0 \leq \varphi \leq 2\pi$  以及  $0 \leq \theta \leq \frac{\pi}{2}$ ）内进行积分，有

$$d\Phi_e = L_e dA \int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin \theta \cos \theta d\theta \quad (\text{A1.2})$$

由于  $\int_0^{\pi/2} \sin \theta \cos \theta d\theta = \frac{1}{2}$ ，因此有

$$d\Phi_e = L_e dA \times 2\pi \times \frac{1}{2} = \pi L_e dA \quad (\text{A1.3})$$

证毕。

## 附录二

### 关于 P112 第 4.3.2.2 小节公式 (4.53) 的证明

为表达简洁起见，本节中分别使用变量  $X$  与  $Y$  对  $CN(\mathbf{g} * \mathbf{M})$  与  $\min[CN(\mathbf{g} * \mathbf{M})]$  中各像素对应的随机变量进行表示。

根据  $\min(\cdot)$  的定义，关于  $Y$  的累积分布函数可表示为

$$\begin{aligned} F(y) &= P(Y \leq y) = 1 - P(Y > y) \\ &= 1 - P[\min(X_1, X_2, \dots, X_K) > y] \end{aligned} \quad (\text{A2.1})$$

式中  $X$  的下标代表通道索引， $K$  表示特征图  $CN(\mathbf{g} * \mathbf{M})$  的总通道数。

由于  $CN(\mathbf{g} * \mathbf{M})$  中各像素在不同通道上的数值可认为满足独立同分布，因此式(A2.1)可改写为

$$\begin{aligned} F(y) &= 1 - P(X_1 > y)P(X_2 > y)\cdots P(X_K > y) \\ &= 1 - P(X_1 > y)^K \end{aligned} \quad (\text{A2.2})$$

根据式 (4.50) 中关于通道归一化操作的定义，我们可近似认为随机变量  $X$  满足标准正态分布，因此有

$$P(X_1 > y) = 1 - P(X_1 \leq y) = 1 - \Phi(y; 0, 1) \quad (\text{A2.3})$$

式中  $\Phi(y; 0, 1)$  表示标准正态分布的累积分布函数。

联立式(A2.2)与(A2.3)有

$$F(y) = 1 - P(X_1 > y)^K = 1 - [1 - \Phi(y; 0, 1)]^K \quad (\text{A2.4})$$

综上， $\min[CN(\mathbf{g} * \mathbf{M})]$  各像素值的数学期望可表示为

$$\begin{aligned} E\{\min[CN(\mathbf{g} * \mathbf{M})]\} &= E(Y) \\ &= \int_{-\infty}^{\infty} yf(y)dy = \int_{-\infty}^{\infty} yF'(y)dy \\ &= K \int_{-\infty}^{\infty} y\varphi(y; 0, 1)[1 - \Phi(y; 0, 1)]^{K-1} dy \end{aligned} \quad (\text{A2.5})$$

式中  $\varphi(y; 0, 1)$  表示标准正态分布的概率密度函数。

证毕。